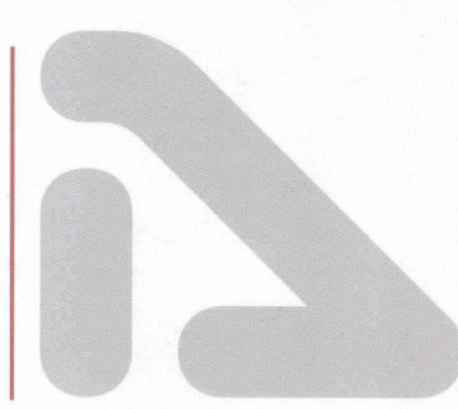
万卷方法

TUJIE AMOS
ZAI YUESHU
YANJIUZHONG
DE
YINGYONG



图解AMOS
在学术研究中的应用

李茂能 著



本书面向所有心理及行为科学领域的学者，帮助他们更加得心应手地运用AMOS。

本书立足AMOS6.0之后到最新版本的新增功能，分别介绍了Bootstrapping抽样、Bayesian（贝叶斯）SEM、类别变量编码法、线性与非线性成长曲线分析、复核效度分析、中介变量与调节变量分析、反映性指标与形成性指标编制的原理、应用及研究实例，并演示如何编写计算程序。

本书特色：

- 反映了SEM的最新动向，如Tetrad SEM、Krippendorff α、幽灵变量法、调节型中介效果、片段线性成长模型等。
- 强调如何针对不同的研究目的、变量类型进行研究方法的选择，对于每一种统计方法在AMOS中的各个步骤，均有研究实例示范和操作界面截图，使读者很容易理解和掌握。
- 详细讲解交互作用效果和间接效果的检验。
- 对于测量工具的编制——量表发展和指标构建，荟萃各家观点，从原理到操作，做了比较详细的介绍。
- 每章均配有习题，方便读者通过练习更好地理解和掌握相关知识点。

本书配套数据及资料可以通过以下途径获取：

访问“万卷方法与学术规范”新浪博客圈

http://q.blog.sina.com.cn/fafang

根据“最新万卷方法图书配套数据（全）下载地址”提示自行下载

上架建议：社会科学

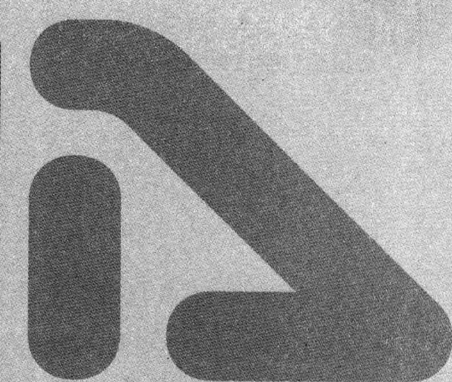
ISBN 978-7-5624-6223-1



定价：39.00元

万卷方法

TUJIE AMOS
ZAI XUESHU
YANJIUZHONG
DE
YINGYONG



图解AMOS
在学术研究中的应用

李茂能 著

重庆大学出版社

图解 AMOS 在学术研究中的应用。原书由台湾五南图书出版股份有限公司出版。原书版权属台湾五南图书出版股份有限公司。
本书简体中文版专有出版权由台湾五南图书出版股份有限公司授权重庆大学出版社在大陆地区出版发行简体字版本。未经出版者书面许可,不得以任何形式复制。
版贸渝核字(2010)第 180 号

图书在版编目(CIP)数据

图解 AMOS 在学术研究中的应用/李茂能著. —重庆:
重庆大学出版社,2011.8
(万卷方法)
ISBN 978-7-5624-6223-1

Ⅰ.①图… Ⅱ.①李… Ⅲ.①统计分析—统计程序,
AMOS—应用—社会科学—研究方法 Ⅳ.①C32

中国版本图书馆 CIP 数据核字(2011)第 141531 号

图解 AMOS 在学术研究中的应用
李茂能 著
责任编辑:林佳木 版式设计:林佳木
责任校对:陈 力 责任印制:赵 晟
*
重庆大学出版社出版发行
出版人:邓晓益
社址:重庆市沙坪坝正街 174 号重庆大学(A 区)内
邮编:400030
电话:(023) 65102378 65105781
传真:(023) 65103686 65105565
网址:http://www.cqup.com.cn
邮箱:fxk@cqup.com.cn (营销中心)
全国新华书店经销
重庆升光电力印务有限公司印刷
*
开本:787×1092 1/16 印张:18.75 字数:399 千
2011 年 8 月第 1 版 2011 年 8 月第 1 次印刷
印数:1—4 000
ISBN 978-7-5624-6223-1 定价:39.00 元

序　言

自从结构方程模型软件 Amos 5.0 版发行之后，Amos 软件除了继续新增统计分析功能之外，还在操作界面、程式语法与操作方法上，出现了一些重大变动。因此，本书前面三章，分别在说明 Amos 5.0/ 6.0 不相容及变动之处并介绍 Amos 6.0/ 7.0 之新增功能与特色。尤其是其新增的贝氏(Bayesian) SEM，将 SEM 推向另一应用新境界。Amos 利用马可夫链蒙地卡罗(MCMC)估计法，展现 SEM 的新风华：适用于小样本、可检验自订参数的假设、利用先验分布解决模式中负的方差估计值的问题，以及使用贝氏插补法，建立多个插补数据集以有效处理遗漏值问题。

第四章之后，则调整了新近 SEM 学术的新知及量表信、效度检验的新技术。这些章节内容主在扩增 SEM 知能：包含 Bootstrapping 在 SEM 上的应用、贝氏 SEM 与 Bootstrapping 的相对效能分析、复核效度之应用、类别变量编码法与回归分析及 SEM 分析之关系、交互作用与间接效果之分析、线性与非线性成长曲线分析、量表发展与指标建构的信度和效度检验。除了理论之深入浅出的简单说明外，大半通过 Amos 软件的操作与学术研究上的实例，“从图中学，从做中学”，以消弭理论与实务上之差距。文中，亦分享了 SEM 学术的国际新动态：例如，Tetrad SEM，Krippendorff a，Ralerting-CV，Taras DM，幽灵变量法、调节型中介效果、片段线性成长模式等。

本书系笔者 2006 年《结构方程模型软件 Amos 之简介及其在测验编制上之应用：Graphics & Basic》一书的延伸与进阶课程，两书具有互补功能，对于 Amos 的操作与 SEM 的运用，更加的深入、充实与完整。近年来我国量化研究的量与质均有显著的提升，如欲更上层楼，端赖学科领域知识的再创新与测量品质的再突破。深信本书对学术创新与研究品质提升，能扮演关键性的辅助角色：可以快速扩增您的学术视野及竞争力，而与国际顶尖的 SEM 学术脉动同步。因此，本书不仅是 Amos 爱好者，亦是 LISREL 使用者的随身宝典。

本书的资料搜集、阅读及撰写费时三年，过程中深感眼力与体力已青春不在，但盼以个人有限的生命换取无限的学术价值。在本书完成之际，要特别感谢过去大学及研究所提携我学术成长的许许多多国内外师长，尤其台湾师大郭生玉老师的即时提拔与教诲，方能奠定日后作学问的稳固基础；而国

外 University of Georgia 教育心理系的Dr. Huberty,Dr. Olejnik 与心理系的 Dr. Lautenschlager 等教授在研究方法学、应用统计学与测量专业上之启迪与要求,更是笔者继志述事的专业标竿。此外,亦要感谢正在美国 UCLA 河滨校区攻读物理学博士学位的邱显智先生,他的限时资料搜集服务,让我的文献搜寻更有效率;亦要感谢嘉义大学数教所林姿饴与教政所吴丽华同学的原始资料,及其他许许多多研究者在期刊中呈现的论文数据,以作为本书活生生的实例论证,对于理论之应用当可不言而喻。更要感激多年来默默耕耘的内人丽云老师,为了经营一个温馨的家,牺牲自我、教育子女,以利笔者之写作。当然五南图书出版公司,不畏惧景气寒冬慨允付梓以及编辑群的多方协助,本书才得以完美体现于读者面前。

最后,本书之撰写过程无不兢兢业业及殚精竭虑,但恐能力未逮或思虑不周,倘有挂一漏万或疏误之处,亦盼海内外方家同好不吝斧正,并借此书抛砖引玉,期待更多的学者投入 SEM 学术、测量统计学之研究,以突破目前学术研究及量表编制上的困境,期能开展我国量化学术研究之新境界。

李茂能　谨识

于嘉义大学

2009 春

目　录

第一章 Amos 5.0 与Amos 6.0 不相容与变动之处

Amos 5.0 拥有轻松易学的操作界面，近几年来其内涵更加强化，使得 Amos 在 SEM 软件界更为耀眼，而获得更多的研究者青睐。Amos 于 2005 年更新上市之 Amos 6.0，其使用者界面、Basic 程序语法及部分操作方法相比 Amos 5.0 出现一些变动，使用者必须加以留神。为避免使用者在应用上之障碍及不便，特先将 Amos 5.0 与 Amos 6.0 在操作界面、程序语法与操作方法之相异之处，简介如下。

第一节 使用者界面变动之处

Amos 6.0 Graphic 界面的功能表单相比 Amos 5.0 Graphic 界面的功能表单上的命名，出现两处的变动。第一是[View/Set]简化为[View]，第二是[Model-Fit]改名为[Analyze]，请参阅图 1-1 Amos Graphic 6.0 之使用者界面。另外，Amos 6.0 Graphic 界面的功能表单亦增列了增益集[Plugins]表单（从 Tools 表单中独立出来），以提供研究者撰写增益集与 6 个已事先设计好的动态连结库，供研究者点选使用。当然撰写增益集需要有撰写 VB. NET 的基本能力，才能得心应手。研究者如何利用 Amos 内建之编辑器撰写增益集之实例，请参阅本书第二章中第六节之范例说明。

从图 1-1 与图 1-2 的功能表单来看，研究者当会发现从 Amos 6.0 起，已能卷动右侧之路径图框，而在[Analyze]的表单中亦增加了贝氏估计法（Bayesian Estimation）与缺失数据填补（Data Imputation）两个统计新功能。另外，从 Amos 6.0 起研究者须先点选[Plugins]下之[Standardized RMR]动态连结库，等出现空白窗口后再执行统计分析，才能在空白窗口中显现 SRMR 指标值，否则无法显示 SRMR 值。

从 Amos Graphics 6.0 之[Analyze]表单知，Amos 新增贝氏估计法（Bayesian Estimation）与缺失数据填补法（Data Imputation）等统计方法，为其他 SEM 统计软件难能向背。尤其是在 Graphics 模式下，研究者尚可在图 1-3 的 VB 窗口中，编写程序以计算及输出自定之统计量数（例如，新的适配度指标），使得 Amos Graphics 兼具便利与弹性之要求。

图 1-1 Amos Graphic 6.0 之使用者界面

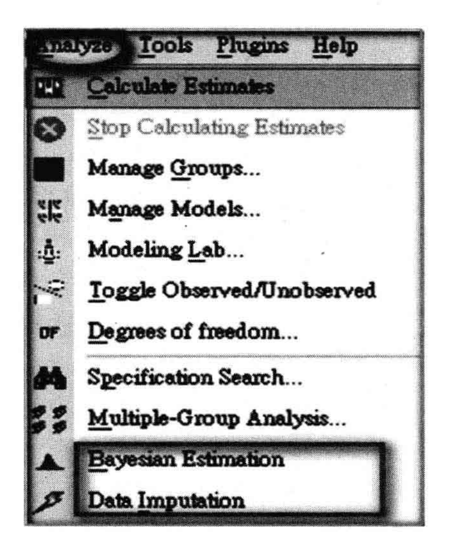


图 1-2 Amos Graphic 6.0[Analyze]表单内容

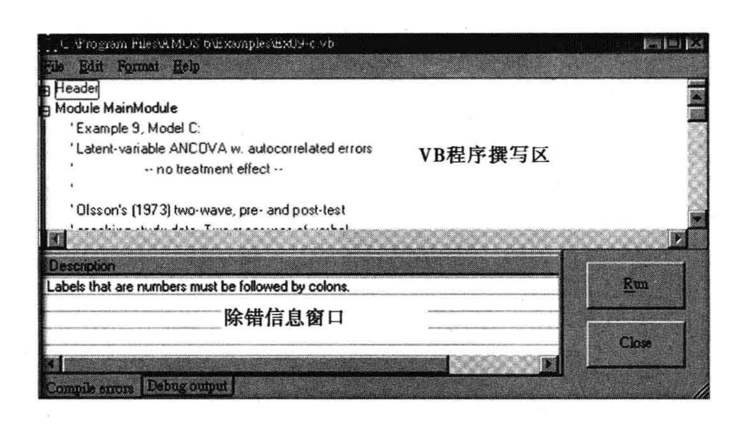


图 1-3　VB 编辑区及除错窗口

第二节　Amos 5.0 与 Amos 6.0 不相容之处

部分 Amos 6.0 所储存之 *.amw 文件(尤其是涉及 Bayesian SEM 的),Amos 5.0 会出现格式不符的警讯而无法正确读入该文件,研究者必须在 Amos 6.0 [File]表单下的[Save As]的对话框中,在存档类型中选取[Amos 5.0 Input File]的文件类型进行存档之后(参见图 1-4),才能再由 Amos 5.0 正确读取。

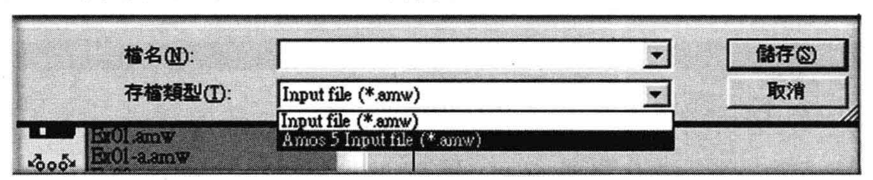


图 1-4　储存 Amos 6.0 之文件格式

另外,在 Amos Basic 语法程序方面,Amos 6.0 已更新到 VB.NET 版本。为了避免与 VB.NET 关键词的冲突,Amos Engine 类别的两个成员:Structure 方法与 Dir 属性,分别更名为 Astructure 与 AmosDir,实例请参见附录一、附录二。另外,TmatrixID enum 与 TMtrixContents enum 的成员名称之字首 ma 均予以取消。例如,maImpliedMoments 须简化为 ImpliedMoments,而 maTotalEffects 须简化为 TotalEffects。

第三节　操作方法相异之处

[Ctrl + B]在 Amos 5.0 中,原系 Tools→Outline 的快速键(Key shortcuts),在 Amos 6.0 中变更为 Analyze→Bayesian 的快速键,但此设定仅限于美式键盘之设定。其次,在 Amos 6.0 的图框中,已无法双击一个对象以打开对象属性窗口,研究者必须改按鼠标右键后,在跳出窗口上点选[Object Properties],再进行对象属性的设定。此外,在 Amos 6.0 界面中,研究者如需网络协助,只须将鼠标移到相关之对象或图像上,按下 F1 之后,即会显示相关的网络协助内容。

第四节 Amos 6.0 之新增四项操作功能

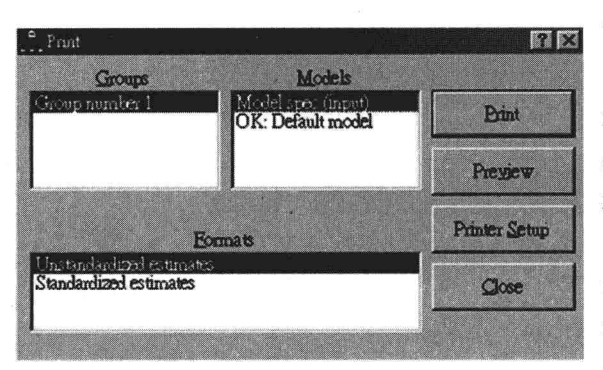


图 1-5 Amos 6.0 打印控制

一、打印路径图前,可以预视

研究者如只须打印路径设计图,请在图 1-5 中 Models 窗口下点选[Model spec],如欲打印路径设计图及相关之未标准化参数估计值,请先在[Formats]窗口中点选[Unstandardized estimates],接着在[Models]窗口中点选[OK: Default model];如欲打印出路径设计图及其标准化参数估计值,则请在[Formats]窗口中点选[Standardized estimates]。

二、改良对象放大与缩小及路径图之卷动

Amos 6.0 使用者将鼠标置于图框中,即可以使用鼠标中间之滚轮放大与缩小路径图,亦可用鼠标上下或左右卷动图框中的路径图(参见图 1-6 之右侧绘图区),使得较大模型的设计更为便捷。

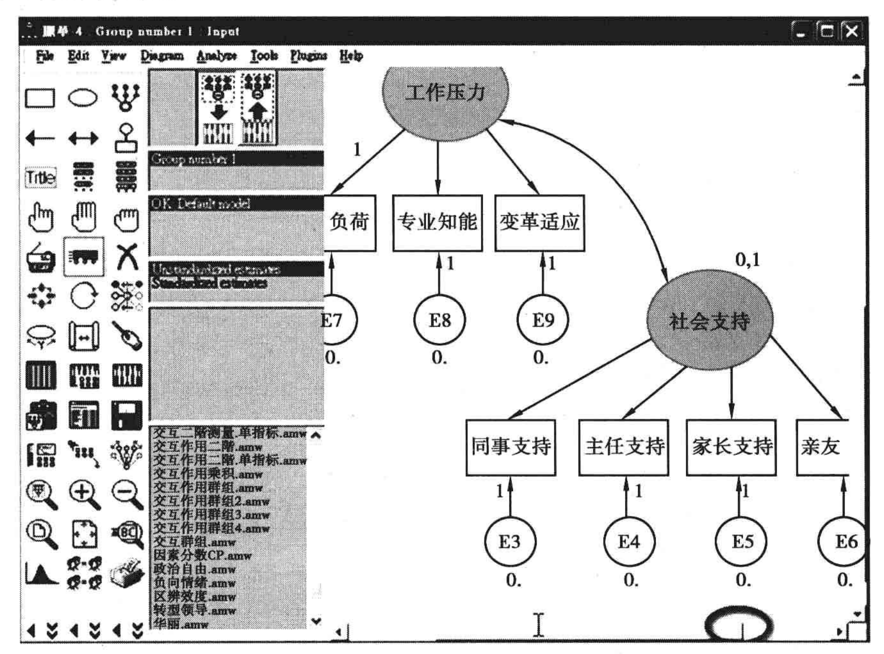


图 1-6 Amos 6.0 之制图区卷轴

三、可以同时开启多个 Amos Graphic 窗口

Amos 6.0 研究者可以同时开启两个以上 Amos Graphic,因此可以同时开启两个以上

不同路径图，以利路径图间之编辑与拷贝。

四、可以拷贝全部或部分路径图

研究者如欲拷贝全部路径图时，可以点选［Edit］功能表单下之［Copy］，或使用［Ctrl + C］拷贝至剪贴簿。如欲拷贝部分路径图时，须先使用事先点选所欲拷贝之对象，再点选 Edit 功能表单下之［Copy］，或使用［Ctrl + C］拷贝至剪贴簿。

本章习题

1. Amos 6.0 与 Amos 5.0 在使用者界面上有哪些重大改变？
2. Amos 6.0 新增了哪些统计功能？

第二章 Amos 6.0 新增功能与特色简介

本章旨在说明 Amos 6.0 新增的功能与特色。Amos 6.0 版的最大特色是提供了贝氏 SEM 估计法、三种新的缺失数据填补法与自定增益集，增添 Amos 不少超越群伦之风华。

第一节 贝氏估计法与马克夫链模拟法

基本上，贝氏估计法乃是结合研究者在数据未搜集前对于测量与结构模型中之待估计参数的先验信念（prior distribution）与实征证据（likelihood of the data），以获取估计参数的事后概率分布（posterior distribution），再进行统计推论。概念上，事后概率分布（$p(\theta|y)$）等于 θ 之事前概率分布乘以观察值 y 之最大概似值：posterior = prior * likelihood。本法特色乃是研究者可以明确地运用对于模式参数的先验知识（如前人之研究或相关理论），以获致参数估计值之事后概率分布，因而可以改善参数估计值及更适合使用于小样本上，并可以避免不合理之模式参数值出现（如负的方差），或进行自订参数函数的估计与检验。Amos 6.0 为了进行总体参数的事后概率分布之推估，会先运用最大概似法初步取得模式参数的估计值，再运用马克夫链模拟法（Markov Chain Monte Carlo，简称 MCMC）模拟模式参数估计值的未确定性。基本上，MCMC 技术具有两种演算型态：（1）Gibbs 漫步法：一次仅从目标条件概率密度函数（target pdf）中抽取一个参数，其余参数保持恒定；（2）Metropolis 漫步法，则所有的参数可能同时产生变动。

Amos 6.0 所采用的 MCMC 演算法为 Metropolis 演算法。首先，从初步的参数估计值中产生许许多多的参数向量，以建立这些相关参数的事后概率分布，此种新 MCMC 模拟法更符合实际及解决更复杂的问题，是贝氏统计的新宠儿。上述这些 MCMC 候选参数向量的产生，是依据 $\theta^{t+1}_{candidate} = \theta^{t} + ax$ 的函数依序产生出来的，其中 x 是正态分布的随机向量（平均数为 0，而协方差矩阵是利用最大概似法所估计来的），而 a 是 MCMC 的微调参数（tuning parameter）。假如前后所产生出来的随机向量相同，Amos 会加以抛弃，继续产生下一个随机向量，一直到所需 MCMC 样本能满足为止。当 Amos 获得研究者所需的样本数之后，会开始进行数据瘦身（thinning），以降低前后样本间之自变相关（autocorrelation），进而降低 MC（Monte Carlo）误差，并提高样本点之代表性。瘦身方法及程序简述如下：Amos 第一次瘦身时会保留偶数样本点，Amos 第二次瘦身时会二中取一，Amos 第三次瘦身时会四中取一，以此类推。因此瘦身三次时，Amos 需要产生 8 个样本才能保留一个样本。

第二节　贝氏 SEM 的 Amos 操作与解释

启动 Amos 贝氏 SEM，有两种方法：

①拉下 Amos［Analyze］表单，点选［Bayesian Estimation］；②点选图像。因为在 Amos 贝氏 SEM 中必须使用原始数据进行统计分析，进行 Bayesian SEM 时，在分析属性窗口中，须勾选［Estimate means & intercepts］。接着，启动 Amos 贝氏 SEM 后，会出现如图 2-1 之 Bayesian SEM 的窗口，Amos 即会开始产生 MCMC 样本。为便利研究者运用图 2-1 中各项功能表单下之功能，Amos 亦均以图像形式呈现其下。因此，使用者可以点选这些图像即可，不必再打开选单点选相关之表单。图 2-1 之底部系事后分布的统计摘要表，表中呈现了事后分布的估计平均数、标准误（SE）、标准差（SD）、聚敛标准（CS）、偏态、峰度、极小值与极大值。其中标准误，系 Monte-Carlo 的事后估计平均数的未确定性指标，当其值 SE 为 0 时即表示完全聚敛（CS = 1.0）。研究者如欲建立参数的置信区间则需使用 SD，才是表示事后平均数与真正参数值的差距。很可惜，在 Bayesian SEM 分析中，Amos 并未提供标准化的参数估计值。如有需要，研究者必须亲自利用相关变量的标准差，计算标准化的参数估计值$\left(= b\,\frac{S_x}{S_y} \right)$。

Bayesian SEM
File　Edit　View　Analyze　Help
1.0000
(500+50,353) * 256　63　29　(500+99,500) * 128
Group number 1

	Mean	S.E.	S.D.	C.S.	Skewness	Kurtosis	Min	Max
Regression weights								
powles71<--71_alienation	1.00	0.00	0.04	1.00	0.18	0.12	0.83	1.2
SEI<--ses	5.22	0.00	0.43	1.00	0.13	0.08	3.49	7.4
71_alienation<--67_alienation	0.61	0.00	0.05	1.00	0.00	0.04	0.41	0.8
71_alienation<--ses	-0.21	0.00	0.05	1.00	-0.14	0.06	-0.44	-0.0
67_alienation<--ses	-0.57	0.00	0.05	1.00	-0.11	0.02	-0.84	-0.3

图 2-1　Amos 6.0［Bayesian SEM］的操作界面

假如，您点选图 2-1 功能表单中的［View］，即可进行各项贝氏统计量之查看与设定，请参见图 2-2 之操作界面。这些表单功能亦出现在图 2-1 上方的功能图像中，它们的功能及操作方法将介绍于后，在此不赘述。换言之，研究者可以不必打开图 2-2［View］下之功能表单，直接点选图 2-1 上方的 6 ~ 11 相关功能之对应图像即可。

为便于使用者操作 Bayesian SEM，特在图 2-1 的操作界面上，针对各图像加注数字编码以利查考，并依编码次序介绍 Bayesian SEM 窗口中各项功能之操作方法如下：

1. 表示打印

打印贝氏统计分析结果。

2. 表示打印预览

预览贝氏统计估计法之统计分析结果。

3. 进行打印设定

4. 进行统计量输出、更新显示、先验概率与 MCMC 参数等之设定，亦可点选图 2-2 中的[Options]功能表单点选，即会出现如图 2-3[Bayesian Options]之交谈窗口。

[Bayesian Options]功能表单下的各项设定简介如下：

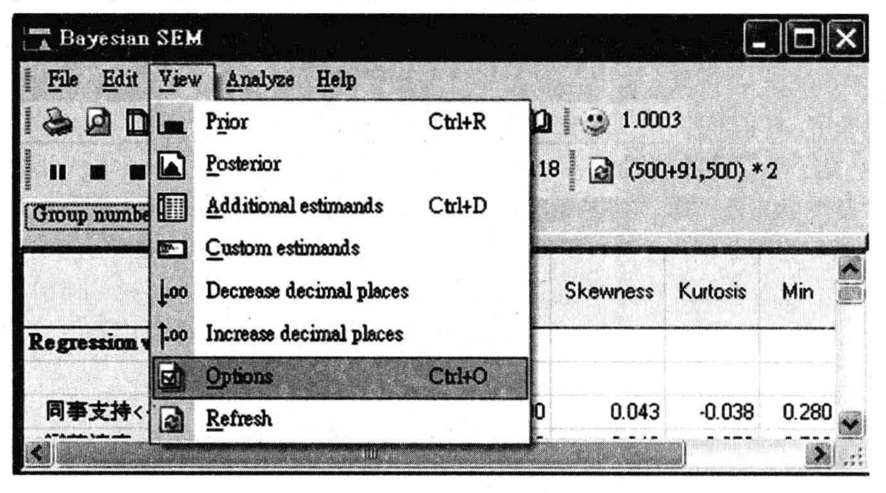


图 2-2　Bayesian SEM 的[View]操作界面

(1)当您点选图 2-3 之[Display]，可以进行各项统计量数之设定。

例如，平均数、标准误、标准差等描述统计量。

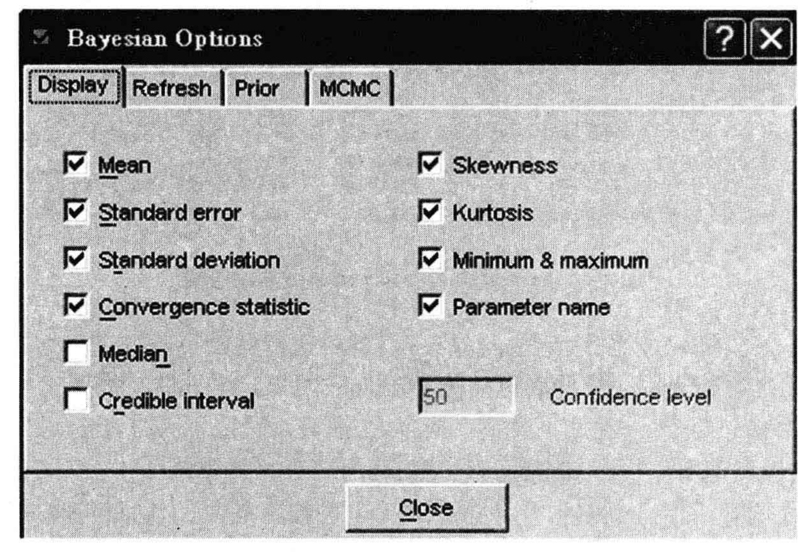


图 2-3　Bayesian SEM 的 Display 操作界面

(2)当您点选图 2-4 之[Refresh]，可以进行抽样结果的统计更新之设定。

例如，您可手动设定，每隔几个观察值或每隔几秒钟就加以更新。

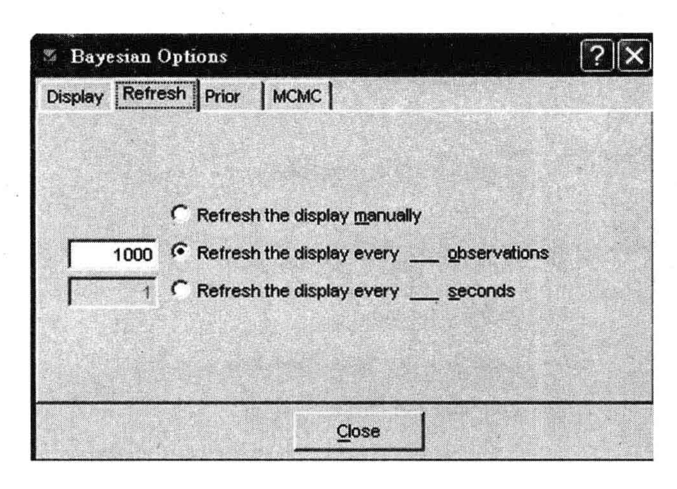


图 2-4　Bayesian SEM 的 Refresh 操作界面

(3)当您点选图 2-5 之[Prior],可以进行参数的 Prior 设定。

在 Prior 表单上勾选[Admissibility test],Amos 会将参数的 prior 设定为0,可以解决非正定矩阵的问题,这与 LISREL Output 中 Ridge 常数的设定功能类似。在 Prior 表单上勾选[Stability test],亦可将参数的 prior 设定为 0,可以解决回归方程解不稳定的问题。

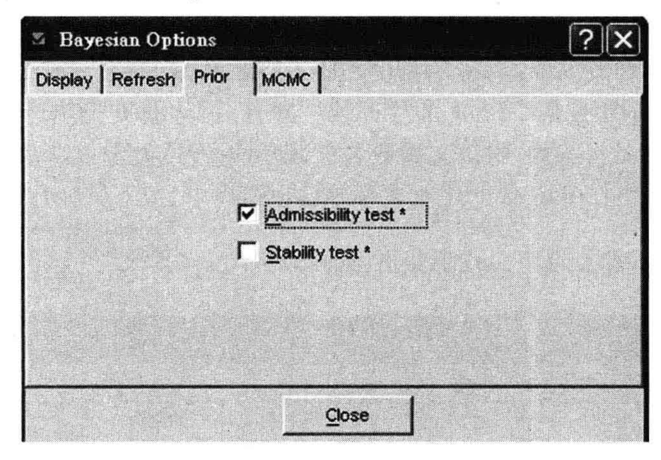


图 2-5　Bayesian SEM 的 Prior 操作界面

(4)当您点选图 2-5 上方之[MCMC]按钮,可以进行各项抽样参数之设定,如图 2-6 左侧窗口所示。

图 2-6 中所示的数据均为内定值,前两个数据之设定,可用来设定将来及目前分析中最大样本保留数(Max observations),当产生的数据点超越此极大值,为了降低前后样本间之自变相关,Amos 会自动瘦身,而可能暂时会导致 CS 值再度升高。因此当您所设定之 CS 标准经过一段长时间仍然无法收敛时,请试着调高此极大值之设定,否则只有降低 CS 标准了。第三个设定系用来设定抛弃样本数(Number of burn-in observations),第四个设定系用来设定微调参数(Tuning parameter),第五个设定系用来设定样本批组数(Number of batches for batch means),计算各批组数据间的自变相关(相关过高时,需要较多的数据产生)及 Monte Carlo 标准误,最后一个方框系用来设定聚敛标准(Convergence criterion),此标准设定愈小,则数据产生的时间通常会愈久。

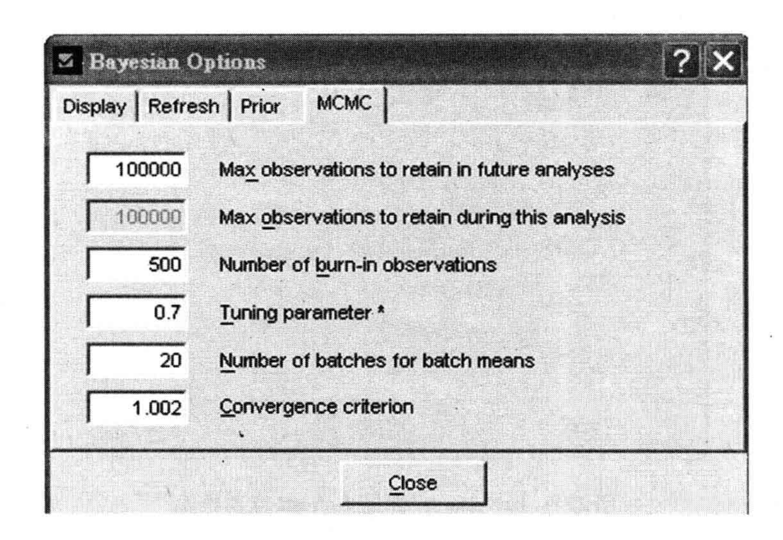


图 2-6 Bayesian SEM 的 MCMC 操作界面

5. 表示拷贝

可拷贝相关参数之事后概率分布的摘要表到剪贴簿中。

6. 观看待估计参数的先验概率分布

须先用鼠标在功能表单[View]中点选 Prior 或图像,就会出现图 2-7 之 Prior 观看窗口,并要求您点选所欲观看 Prior 的估计数。接着,再移向贝氏统计分析摘要表中(如图 2-1)的待观看的估计参数,按下该参数名称,即可查看或设定该参数之先验概率分布。更直接简便的方式是在 Bayesian SEM 之结果输出窗口中,按下鼠标右键点选欲观看之参数列,即会出现要求您点选事后或事前概率分布的小窗口 Show Posterior Show Prior 。

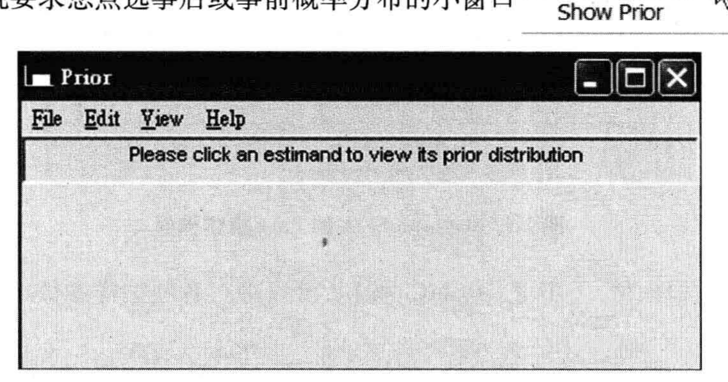


图 2-7 Bayesian SEM 之 Prior 观看窗口

换言之,研究者根据参数在总体中分布的既存知识、实际观察或专家的评估,进行各参数 Prior 统计量数之设定。Amos 中各参数内定的 Prior 为等概率(Uniform)分布,此系均分性的先验概率分布(Diffuse prior distribution,又称为 Non-informative Prior distribution),会导致与传统之 SEM 分析结果相同,因为事后概率分布($p(\theta \mid y)$)等于 θ 之事前概率分布乘以观察值 y 之最大概似值。Prior 为等概率较适合于一般的数据分析,如果研究者的兴趣在预测下一个事件,则最好使用其他合适之先验分布。由图 2-8 知,研

究者除了可以选择等概率 Uniform 或正态 Normal 之外，亦可自定先验概率分布，以提高预测力。

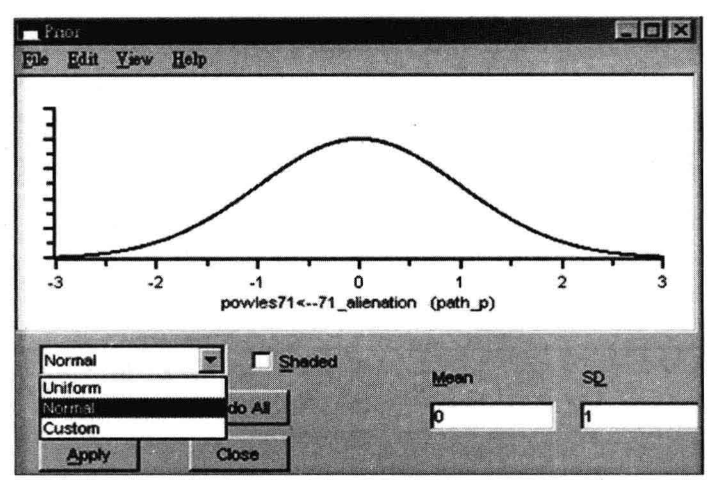


图 2-8 Bayesian SEM 的操作界面：查看及设定参数之先验概率分布

研究者如欲自定“参数之先验概率分布”，须先点选图 2-8 中[Custom]选目，之后就可在图 2-9 中拉着铅笔之图像到所欲更改分布型态之位置，并按下鼠标左键移动铅笔，即可绘制自定参数之分布型态。研究者如遇 path_p 之参数出现不合理之数值(如为负值)，亦可手动将图 2-9 右下角之 Lower bound，设定为 0，亦可解决不适当解(Improper solution)的问题，当然研究者也可使用图 2-5 的设定方法让 Amos 自动解决不适当解的问题。不过研究者对于参数之先验概率分布的设定，应有理论或经验之根据，否则可能导致更大的误差。值得注意的是，随着原始样本之增大，先验概率分布的影响力就逐渐下降。除非您的样本非常小或您的模型或先验概率分布与数据有很显著的矛盾，否则您会发现虽然使用不同的先验概率分布，贝氏估计值不会有很大的更动。

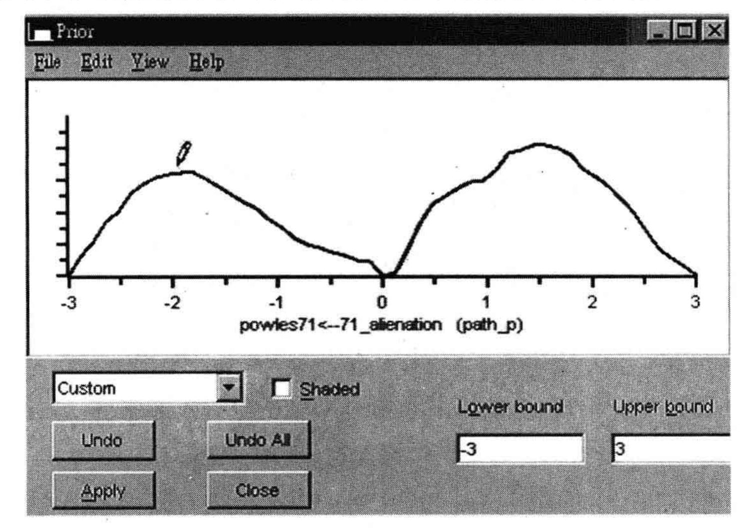


图 2-9 Bayesian SEM 的操作界面：自定参数之先验概率分布

7. 观看待估计参数的事后概率分布

操作方法如事前概率分布,须先用鼠标在功能表单[View]中,点选 Posterior 或图像后,再移向贝氏统计分析摘要表中欲观看的估计参数,按下该参数名称即可查看,或设定该参数之事后概率分布(参见图 2-10)。

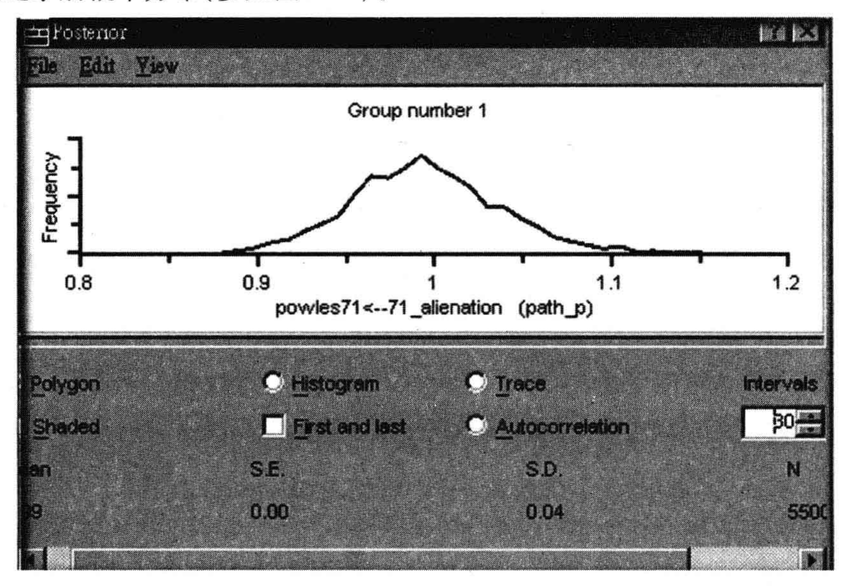


图 2-10　Bayesian SEM 的操作界面:查看参数之事后概率分布

8. 进行额外参数的估计

本功能旨在计算模型参数之外的其他参数的估计值,如直接效果的估计,估计之过程中会出现图 2-11 与图 2-12 之窗口。额外参数的估计,一次仅能显示一种估计值于窗口中。

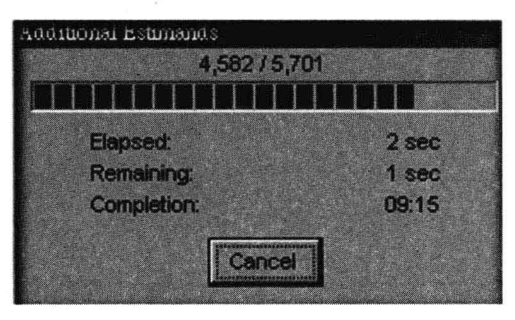


图 2-11　额外参数估计的过程

当这个额外估计窗口出现后,即可显示额外参数的估计结果表单,如图 2-12 所示。

9. 利用 VB. NET 或 C# 撰写模式参数函数,要求 Amos 估计新参数

过去 Amos 可以利用 Bootstrapping 技术,估计直接效果与间接效果之大小、P 值与置信区间。在 Bayesian SEM 中则可使用如图 2-13 之 VB. NET 程序,进行直接效果(c)与间接效果(a * b)间差异分析与统计检验。操作时,研究者须先终止 MCMC 抽样,之后按下[Custom estimates]钮,就会出现图 2-13 之自定参数窗口,供研究者撰写程序。图 2-13

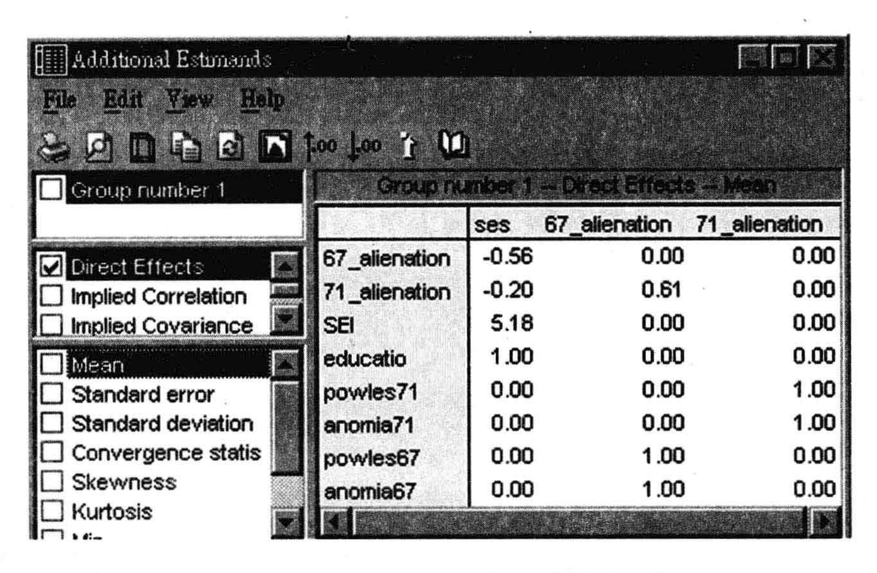


图 2-12　Bayesian SEM 额外参数的估计结果

中第一个副程序(Public Sub),旨在宣告新估计值之名称,下一个估计值计算函数(Public Function),则在撰写新估计值之计算程序码,旨在计算直接效果、间接效果、两者间之差异值及间接效果小于 0 和间接效果比直接效果小的概率。窗口程序中[Implements Iestimand]表示要实作 Estimand 的界面(interface),亦即要实际撰写界面宣告所定义之属性或方法的程序码。本例子系使用 Amos 提供的实例——Example 29 作为示范,文件名为 Ex29. amw。写完程序后,按下[Run]即可在图 2-14 窗口中显现统计分析结果。

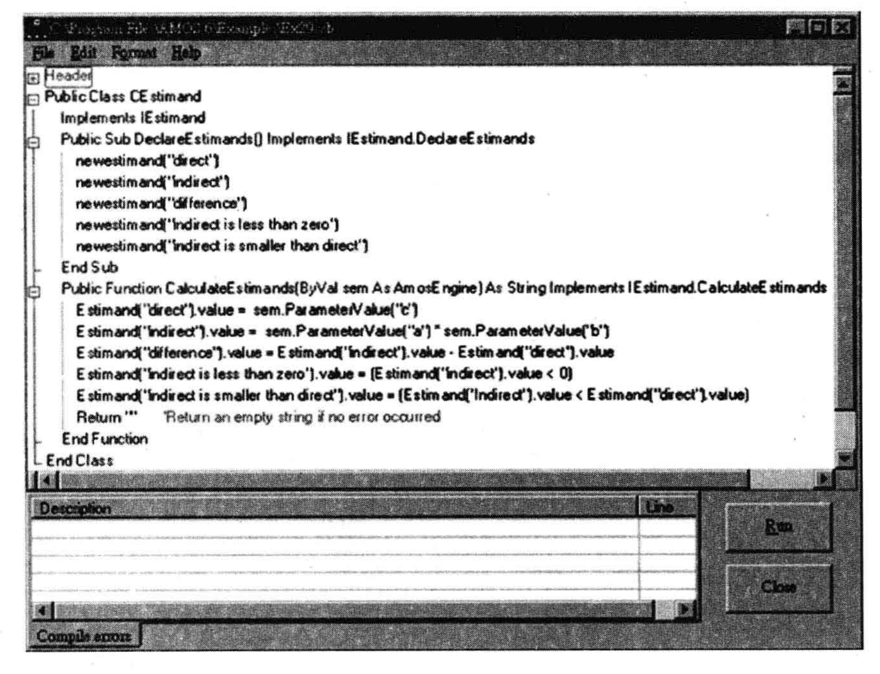


图 2-13　自定参数之统计检验的 VB. NET 程序设计

图 2-14[Dichotomous Estimands]一行中,P1 表示前 1/3 的 MCMC 样本中,该自订估

Custom Estimands

File Edit View Help

Numeric Estimands	Mean	S.E.	S.D.	C.S.	Skewness	Kurtosis	Min	Max
direct	-0.205	0.002	0.052	1.001	-0.153	0.026	-0.409	-0.005
indirect	-0.342	0.002	0.040	1.001	-0.369	0.327	-0.571	-0.228
difference	-0.137	0.003	0.073	1.001	-0.062	0.077	-0.484	0.136

Dichotomous Estimands	P	P1	P2	P3
indirect is less than zero	1.000	1.000	1.000	1.000
indirect is smaller than direct	0.971	0.957	0.976	0.980

图 2-14　Bayesian SEM 三个自订估计量的边缘事后分布

计值（如 indirect is less than zero 或 indirect is smaller than direct）为真之比率（例如，间接效果小于直接效果之比率为 0.957）；P2 表示中间 1/3 的 MCMC 样本中，该自订估计值为真之比率；P3 表示后 1/3 的 MCMC 样本中，该自订估计值为真之比率；而 P 为全体 MCMC 样本中，该自订估计值为真之比率。在本例间接效果小于直接效果之比率为 $p=\frac{\#(\text{indirect}<\text{direct})}{N_{MCMC}}=0.975$，它与 P1、P3、P3 之值亦甚为接近，表示 P 值甚为稳定，为一很好的统计估计值；而间接效果小于零之值为 $p=\frac{\#(\text{indirect}<0)}{N_{MCMC}}=1.000$，因此可以很肯定地说间接效果是负值（x = 0.342）。

10. 增加小数位数

按下此按钮即可逐渐增加小数位数。

11. 减少小数位数

按下此按钮即可逐渐减少小数位数。

12. Amos 版权宣告

13. 线上协助

研究者如需线上协助，可按下此钮寻求线上即时说明。

14. 聚敛统计图示

MCMC 的聚敛是指抽样分布的聚敛，当推论不再依赖最初样本起始估计值时，即表示获得聚敛，亦即我们终于获得一个无偏事后概率分布了，研究者即可停止抽样。聚敛统计量（CS）小于内定 1.002 时（此标准研究者可以自调），Bayesian SEM 窗口会出现笑脸，否则会出现红色的哭脸。

15. 1.0000 聚敛统计量

CS（Convergence satatistics）乃是当 MCMC 抽样数目趋于无限大时，用来衡量降低估计数不确定性之指标。CS 等于 1 表示完美聚敛（perfect convergence），研究者可以不必再产生更多的样本。通常如 CS 大于 1.1，则仍须继续产生更多的 MCMC 样本以降低估计数之不确定性，Amos 采取较保守的内定阈值 1.002。Amos 贝氏 SEM 功能表单上的 CS

值,代表参数中最大之 CS 值。

16. ▮▮ 暂停抽样

按下此键可暂停抽样,再按一次则会恢复抽样。当您认为 MCMC 抽样数达到满意程度时,即可按下此钮结束 MCMC 抽样。

17. ■ 抛弃已累积的样本,重新抽样

18. ■ 关闭 Bayesian SEM 窗口

19. (500+50,353) * 256 样本大小

图 2-1 中(500 + 50353) * 256 表示 Amos 曾产生了 13018368(50853 * 256) 个样本,但只保留了 50353 个样本(式中的 * 256 表示 Amos 已进行了 8 次瘦身的工作(2^8 = 256),因而仅保留了 1/256),式中 500 系内定 burn-in 样本数(抽样分布未聚敛前所抛弃的样本数)。因为初期的样本向量通常不是取样于参数事后概率分布,Amos 会抛弃前 500 个样本,仅利用剩下来的 50353 个样本进行参数之事后概率的推估。

20. 63 每秒 MCMC 所产生的样本观察值数

21. .29 样本接纳比率

MCMC 要能有效运作,样本接纳比率最好在0.2 ~0.5。假如样本接纳比率低于0.2,可调低微调参数(内定值为0.7),假如样本接纳比率高于0.5,可调高微调参数。

22. 更新抽样结果

按下此按钮可以更新参数之事后概率分布摘要表及分布图,以反映最新保留样本的统计结果。

23. (500+99,500) * 128 系指前次样本更新之大小

Amos 在产生 1000(内定)个样本后,才会更新结果。

(500 + 99500) * 128 中的 * 128 表示 Amos 已在进行瘦身的工作,128 表示已瘦身 7 次(2^7 = 128)。

第三节 MCMC 聚敛之诊断分析

Amos 提供 CS 统计量与图表分析两种聚敛分析之诊断方法。读者应先分辨此两种 MCMC 聚敛的意义。

第一种聚敛称为抽样分布之聚敛(Convergence in distribution),意指分析之样本的确是从参数的联合事后分布中抽取出来。由于起始样本点可能无法真正代表事后分布,Amos 内定抛弃起始的 500 个样本点,这是保守的做法,大部分的问题可能不需这么多。

第二种聚敛称为事后摘要统计之聚敛(Convergence of posterior summaries),意指分析之样本是否够大而能获得正确的事后摘要统计量(如 posterior mean)。CS 等于 1 表示完美聚敛(perfect convergence),更多的样本亦无法再提高统计量之精确性。因此,当 CS 逼近于 1.0 时研究者就可终止样本的产生工作。通常如 CS 大于 1.1 时,则仍须继续产生更多的 MCMC 样本以降低估计数之不确定性,Amos 采取较保守的内定阀值 1.002。

除了前述之 CS 值可供诊断贝氏 MCMC 方法所得之统计量是否聚敛之外,Amos 还提

供以下几种图示法，协助研究者判断抽样分布是否已收敛：前 1/3 与后 1/3 分布一致性分析图（参看图 2-15 和图 2-16）、轨迹图（或称时间系列图，参看图 2-17 和图 2-18）与自变相关图（参看图 2-19 和图 2-20）。

由图 2-15 知，抽样分布在未收敛前，前 1/3 与后 1/3 抽样分布之一致性，具有明显之差异。

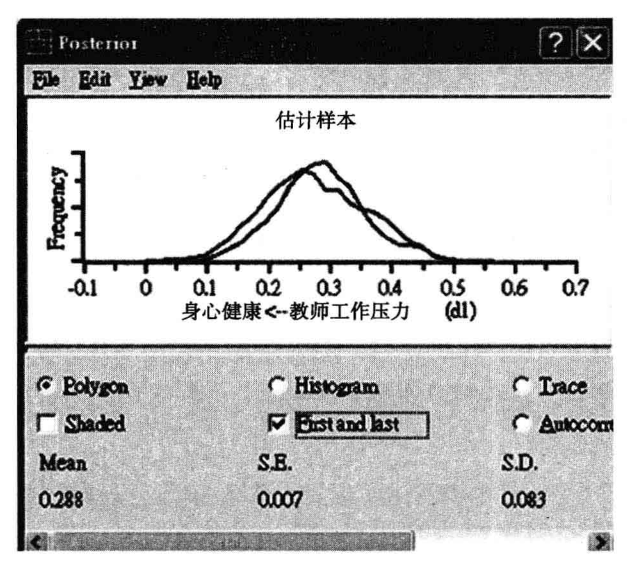


图 2-15 前 1/3 与后 1/3 分布一致性分析图：收敛前

由图 2-16 知，当前 1/3 与后 1/3 分布曲线趋于一致时，表示该参数之事后概率分布已接近收敛。

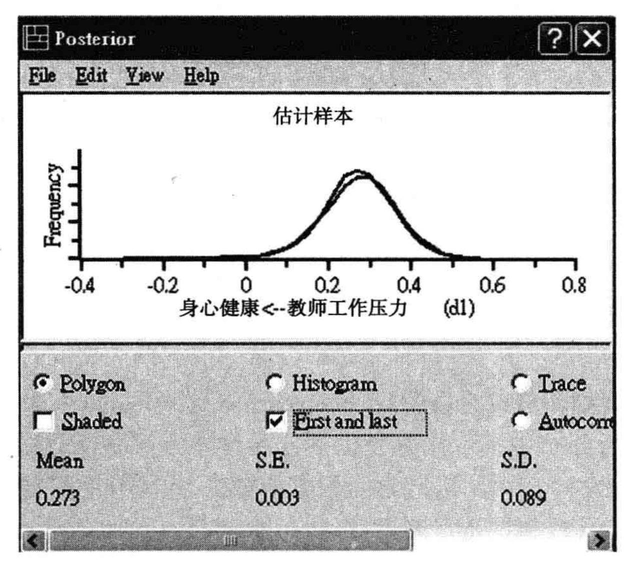


图 2-16 前 1/3 与后 1/3 分布一致性分析图：收敛后

由图 2-17 知，参数轨迹图可用来评估 MCMC 抽样分布是否已收敛，未收敛前该参数值之抽样分布会有明显的上下震动或随意漂流的现象，但在图 2-18 中，则无此现象。

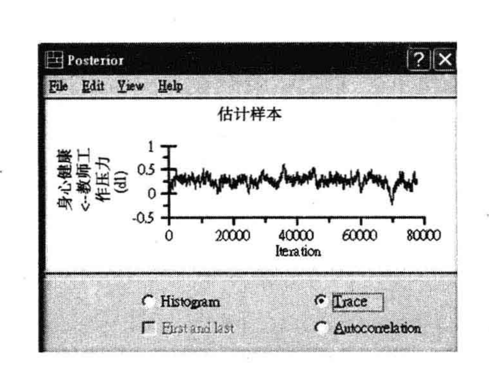




图 2-17　Bayesian SEM 参数轨迹图:收敛前

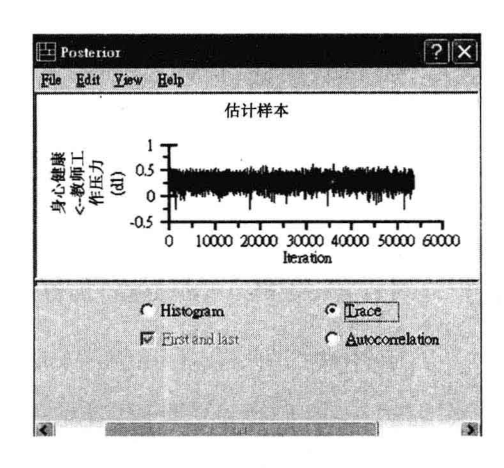


图 2-18　Bayesian SEM 参数轨迹图:收敛后

由图 2-19 知,此参数之自变相关在 Lag 100 时(回圈间距相差 100 的两两抽出样本点间的相关),其自变相关仍未趋于 0。

由图 2-20 知,自变相关在 Lag 50(回圈间距相差 50 的两两抽出样本点间的相关)之后,自变相关已接近于 0,表示 MCMC 抽样已脱离起始点,其抽样分布已达收敛。

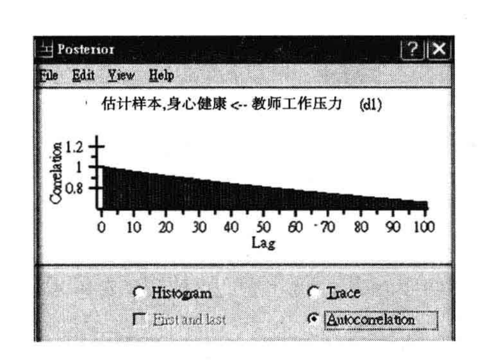


图 2-19　自变相关图:收敛前
(Lag:样本点出现之回圈间距)

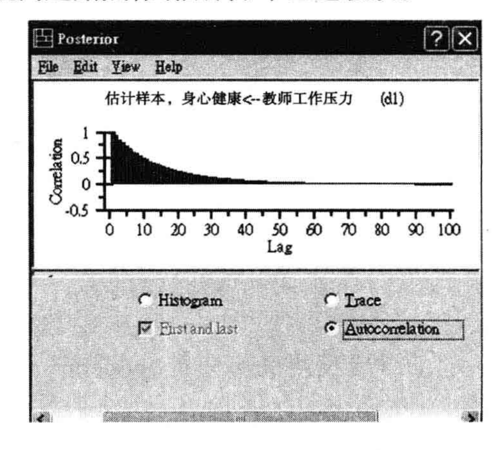


图 2-20　Bayesian SEM 自变相关图:收敛后

第四节　缺失数据填补法

由于缺失数据的删除法(如 Listwise 或 Pairwise 删除法)常会因样本变小而导致抽样误差增大、非正定矩阵、流失有用信息、降低统计检验力及削弱参数估计值之正确性。因此,目前 SEM 使用者倾向于使用数据填补法。因应时势所需,Amos 6.0 提供了三种缺失数据填补(Data imputation)方法(参见图 2-22),以填补缺失值:

一、回归填补法(regression imputation)

本法乃是使用个案之非缺失值,去预测该个案之缺失值的数据,本填补法是单一固定值填补法。Amos 在取得最大概似估计值之后,会将它设定为模型之参数值,再利用线性回归分析进行缺失值的预测。

二、随机回归填补法(stochastic regression imputation)

本法与回归填补法类似,除了使用观察数据之最大概似值取得回归预测值后,一般常加上残差后再填补缺失值(例如,SPSS 的回归填补法即可采用此设定)。不过,Amos 在本填补法中,乃先根据实得观察值建立缺失值的条件概率分布,再从缺失值的条件概率分布中随机抽取填补值,如有需要可产生多个填补过的完整数据集,可利用图 2-22 窗口中[Number of completed datasets]加以设定。

三、贝氏填补法(Bayesian imputation)

本法与随机回归填补法类似,但视未知的模型参数为随机估计值而有一个相对应之概率分布(a probability distribution),而非固定之最大概似估计值。在 Amos 中的贝氏填补法是运用多元填补技术(Multiple imputation)与事后模拟(Posterior simulation),每一缺失值都会使用多个(m > 1)模拟值取代之,以估计这 m 个数据集之平均数、变异量与相关之 p 值或置信区间。

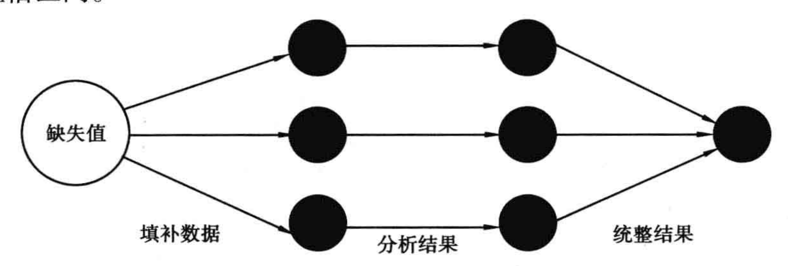


图 2-21 缺失数据多元填补过程

由图 2-21 知,多元填补技术涉及三个步骤:

①每一缺失数据点以 m 个(>1)数据点填补,这些数据点系通过 Markov 链技术,从缺失值的事后预测概率分布(posterior predictive distribution)中,随机抽取 m 个独立观察值;

②利用统计分析软件如 SAS、SPSS、Amos 等分析 m 个数据集(一般来说,5 ~ 10 个已足够);

③整合 m 个分析的结果(具体统计分析方法,请参看下节简介及 Amos 6.0 操作手册第 479 页之说明)。

因而,含有缺失值的贝氏数据分析,涉及以下两大工作:多元填补技术与事后模拟。前者,系从缺失值的事后预测概率分布(posterior predictive distribution)中,随机抽取 m 个独立观察值;后者,系于缺失数据填补后,从模型参数的事后概率分布中进行随机抽样。研究者如有需要可产生多个填补过的完整数据集,设定方法亦如前述随机回归填补法。这些填补过的完整数据集,如欲放在一个文件中,请在图 2-22 中点选[Single output file],否则请点选[Multiple output files],将各个数据集存放于不同名称的文件中。如将各个数据集存放于同一个名称的文件中,将来欲个别分析每一数据集时,可利用 SPSS 的 Split file 功能进行文件分割后,再进行个别数据集的统计分析。

第五节　Amos 的数据填补操作步骤

在 Amos 中进行数据之填补，需要点选 Amos 界面[Analyze]下之[Data Imputation]，即会出现如图 2-22 之设定窗口，选择您所欲采用的数据填补方法。图 2-22 所点选的填补方法是贝氏填补法(Bayesian imputation)，按下底部[Options]按钮，即会出现本图之右侧[Options]窗口，进行贝氏填补法演算法的参数设定：观察值数目、自变相关之最大值、微调参数。

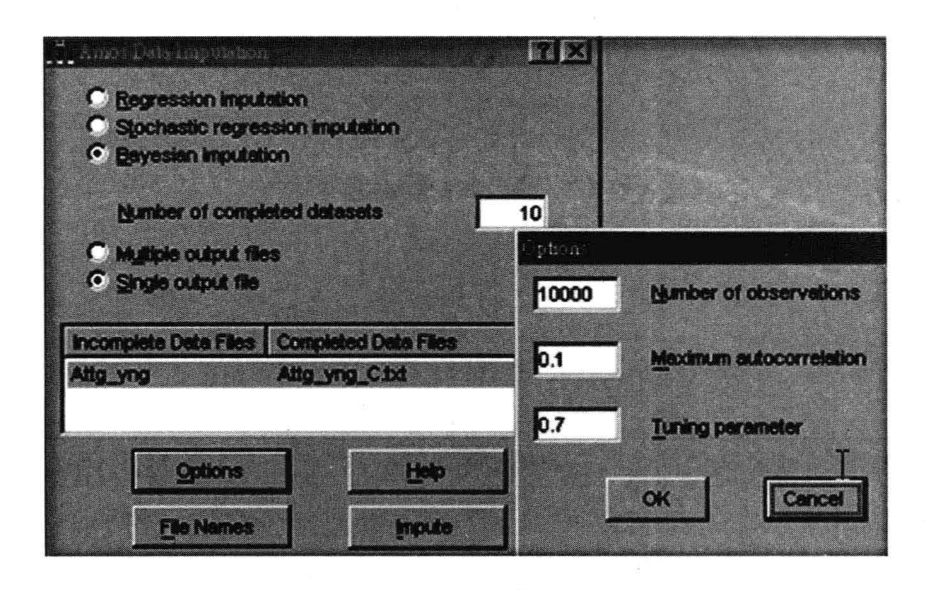


图 2-22　Amos 缺失数据填补法设定窗口

在 Amos 中实际进行贝氏填补法时，须进行 MCMC 自动产生观察值数目(100 ~ 30000)、参数间之自变相关(maximum autocorrelation, 0 ~ 1.0)及微调参数(tuning parameter, 0 ~ 1.0)等设定，参见图 2-22 之右侧[Options]窗口。当 MCMC 观察值达到所设定之数目后，如每一参数之自变相关低于设定值时(如 0.1)，样本之抽样才会终止，否则会抛弃奇数样本点后继续抽样。Markov 样本链之自变相关过高，意味着样本链锁定在事后概率分布的某一区域内，会导致不佳变异量估计值。瘦身细节如同前述，因而 Amos 会自动进行 Markov 样本链的瘦身，以降低样本链的相关性。至于微调参数之设定，旨在控制前后参数估计值之变化量，微调参数愈大，前后参数估计值之变化量则愈大(内定值为 0.7)。所有相关设定完后，按图 2-22 中之[Impute]按钮，Amos 即会执行数据填补与存档之工作，存档之文件名请用图 2-22 中之[File Names]按钮进行设定，以便日后统计分析使用。

在此值得一提的是，Amos 在进行随机回归填补法、贝氏填补法与贝氏 SEM 时，会大量使用到随机数(random numbers)，而此随机数的选取须视随机数的起始值(seed)而定。Amos 的内定值是每一次的样本点抽样，均使用不同的随机数起始值(每次递增 1，Increment the current seed by 1)，研究者如欲获得相同的样本数据集(可复制前面的数据分析结果)，必须使用相同的随机数起始值。操作时请在 Tools 功能表单下的 Seed

Manager 窗口中(参见图 2-23),点选[Always use the same seed]。

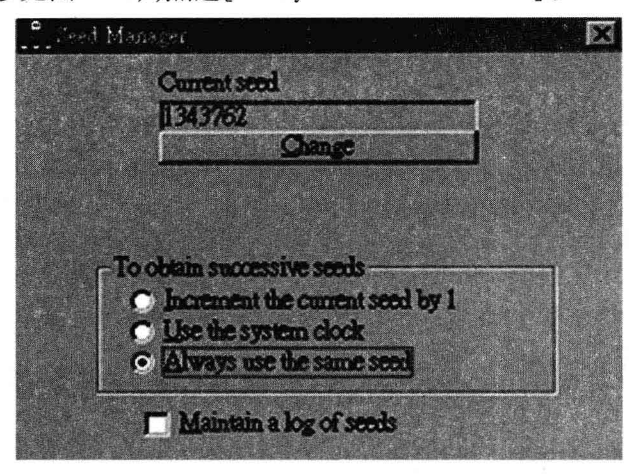


图 2-23 随机数起始值设定窗口

至于,如何整合 m 个个别数据集的统计分析结果,过去常依靠研究者之笔算,随书所附之 MI 统计分析副程序(操作方法及步骤请参见附录五(SEMCAI. xla),可免去研究者手算之苦。兹以 Amos 7.0 操作手册第 475 页的实例,具体说明如下。假设有 m = 10 个回归分析的结果,其各个数据集的回归系数估计值与标准误详如表 2-1:

表 2-1 10 个个别数据集的统计分析结果

Imputation	ML Estimate	ML Standard Error
1	1.106	0.160
2	1.080	0.160
3	1.118	0.151
4	1.273	0.155
5	1.102	0.154
6	1.286	0.152
7	1.121	0.139
8	1.283	0.140
9	1.270	0.156
10	1.081	0.157

今令 $\hat{Q}^{(1)}$ 为样本 t 的回归估计值,因此 $\hat{Q}^{(1)} = 1.106$, $\hat{Q}^{(2)} = 1.080$,其余以此类推。又令 $\sqrt{U^{(t)}}$ 为样本 t 的回归估计值的标准误。因此,$\sqrt{U^{(1)}} = 0.160$, $\sqrt{U^{(2)}} = 0.160$,其余以此类推。那么多元填补回归系数之估计值(the multiple-imputation estimate of the regression weight),等于这 10 个数据集个别回归系数估计值的平均数。

$$\overline{Q} = \frac{1}{m}\sum_{t=1}^{m} \hat{Q}^{(t)} = 1.172$$

至于多元填补的回归系数参数估计值之并组标准误的步骤,演示如下:

(一)计算组内填补方差的平均数

$$\bar{U}=\frac{1}{m}\sum_{i=1}^{m}U^{(t)}=0.0233$$

(二)计算组间填补方差

$$B=\frac{1}{m-1}\sum_{i=1}^{m}(\hat{Q}^{(t)}-\bar{Q})^{2}=0.0085$$

(三)计算总方差

$$T=\bar{U}+\left(1+\frac{1}{m}\right)B=0.0233+\left(1+\frac{1}{10}\right)0.0085=0.0326$$

(四)计算多群组标准误

$$\sqrt{T}=\sqrt{0.0326}=0.1807$$

(五)检验回归系数等于0

$$\frac{\bar{Q}}{\sqrt{T}}=\frac{1.172}{0.1807}=6.49$$

本统计量为 t 统计量,其自由度计算公式为:

$$v=(m-1)\left[1+\frac{\bar{U}}{\left(1+\frac{1}{m}\right)B}\right]^{2}=(10-1)\left[1+\frac{0.0233}{\left(1+\frac{1}{10}\right)0.0085}\right]^{2}=109$$

上述相关之统计量与其置信区间,除了使用附录五笔者所设计之 SEMCAI 软件之外,尚可以利用 Schafer(2000)的 NORM2003. exe 计算器快速取得。其免费下载网站是:http://www. stat. psu. edu/ ~kls/misoftwa. html。其操作的程序简单说明如下:

(一)准备原始数据并依图 2-24 格式建立文件

回归系数在前,相关之标准误在后,接着存入纯文字文件如 Reg. dat。

```
檔案(F) 編輯(E) 格式(O) 檢視(V) 說明(H)
1.106  .16
1.080  .16
1.118  .151
1.273  .155
1.102  .154
1.286  .152
1.121  .139
1.283  .14
1.270  .156
1.081  .157
```

图 2-24 原始数据文件格式

（二）打开图 2-25 中 Norm 程序的[Analyze]表单，点选[MI inference：Scalar]

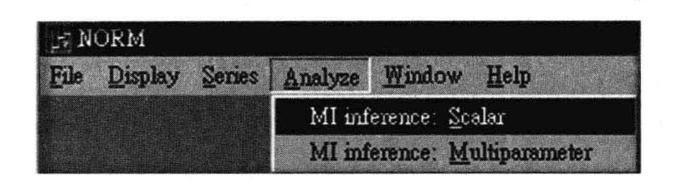


图 2-25 Norm 程序的[Analyze]表单

当您点选[MI in ference：Scalar]之后，会出现图 2-26 之文件选用窗口。

（三）呼叫资料文件 Reg. dat

按下图 2-26 中之[开启]钮之后，会出现图 2-27 之设定窗口。

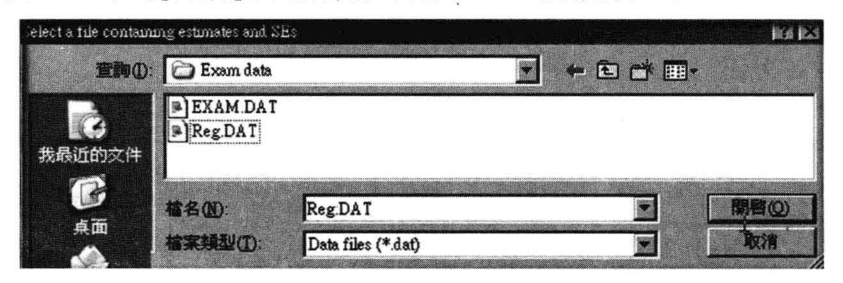


图 2-26 Norm 程序所用数据文件的选用表单

（四）在图 2-27 窗口中键入相关之资讯

在图 2-27 中键入输出文件位置及名称、标题名称、估计值文件位置、建文件格式（本例为 stacked columns）、估计参数个数（Number of estimands）、填补数据集的个数（Number of imputations）后，并键入参数估计值的置信水平（Confidence level），即可按下[Run]执行之。

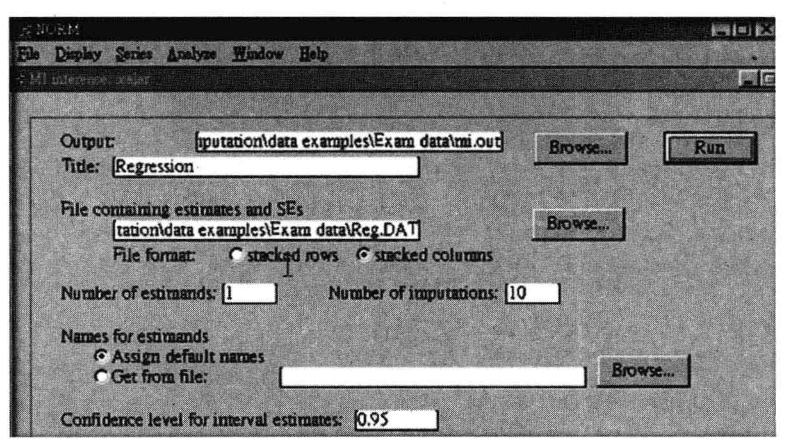


图 2-27 NORM 程序计算 MI 平均值及参数检验的设定表单

执行后,分析结果详如图 2-28 所示。

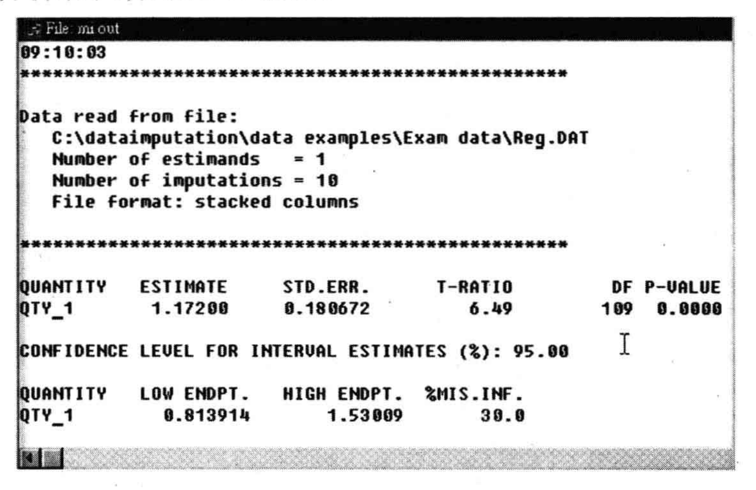


图 2-28　NORM 程序的 MI 输出报表

(五)跑出的 MI 统计结果

由图 2-28 知,NORM 程序所得出的 MI 相关之统计量与前述手算出来的统计量几乎完全一样,并且能提供 t 检验的 p 值与置信区间。

第六节　利用 Amos 内建之编辑器撰写增益集:以 SRMR 为例

为便利研究者自订 SEM 统计参数,Amos 允许研究者,运用其内建之程序编辑器撰写增益集(Plugins)。兹将 Amos 增益集之撰写与执行步骤,简述如下:

一、撰写增益集之步骤

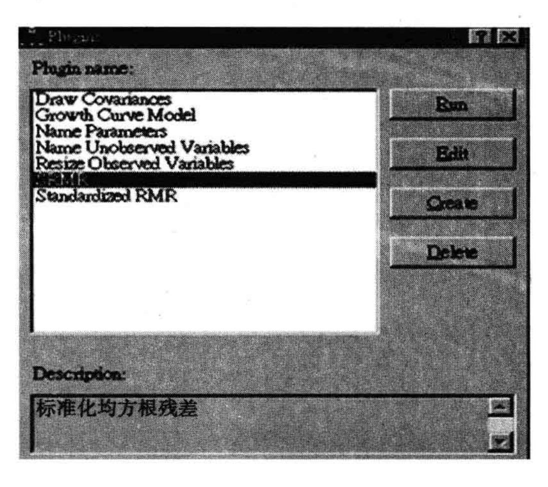


图 2-29　Plugins 撰写操作画面

1. 拉下第一章图 1-1 中 Amos 功能表单[Plugins]选单,点选 Plugins 选目,开启如图 2-29 之画面。

2. 点选[Create]按钮,在 Amos 内建之 VB 编辑器内(参见图 2-30),撰写相关程序,至于完整之程序语法范例,请参考下文所述。

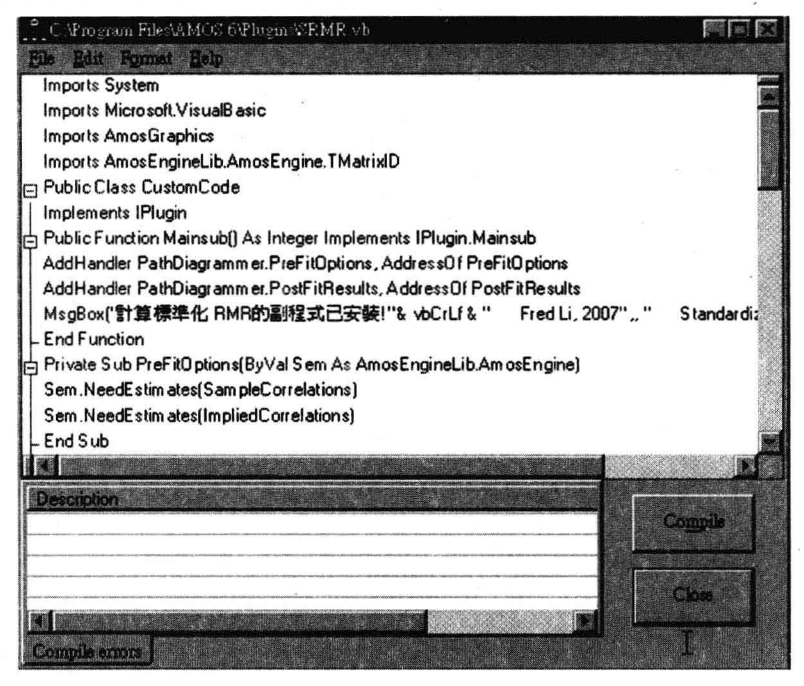


图 2-30　Amos 内建程序编辑器

3. 按图 2-30 中[Compile]按钮,进行 VB 程序之编译工作,假如有语法上之错误,会在 Description 窗口中出现相关之警讯,以利除错。没有任何错误时,即可打开[File]选单进行 VB 程序存档,最好存在 Amos 之 Plugins 目录夹中。

4. 按下图 2-30 中[Close]关闭 VB 编辑窗口。

二、执行自定增益集之步骤

1. 开启如图 2-29 之 Plugin name 窗口,在其窗口内点选[SRMR]。

2. 按下图 2-29 中之[Run]执行钮。

3. 接着,出现如图 2-31 的 SRMR 增益集已安装之讯息框,按下图中之[确定]钮,表示您已可在接下来的统计分析时使用它。

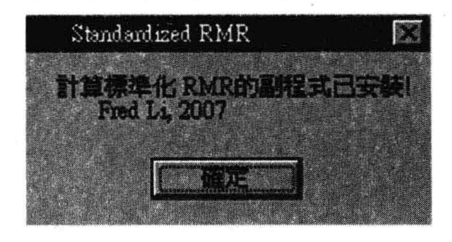


图 2-31　SRMR 之安装讯息窗

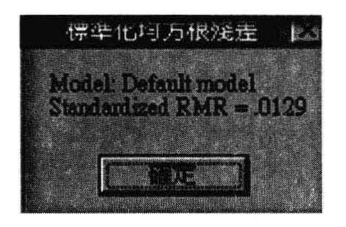


图 2-32　SRMR 之显示窗口

4. 之后执行统计分析，在未关闭 Amos 之前都会自动显示 SRMR 之适配度指标，如图 2-32 所示。换言之，在新启动的 Amos 之后均须重新启动新的自定增益集，才能执行额外之 SRMR 统计分析。

三、SRMR 增益集范例程序内容

研究者可于图 2-30 中之 Amos 内建之编辑器中，撰写增益集程序码，以下系 SRMR 之完整程序码内容，供研究者参考应用：

```
Imports System
Imports Microsoft. VisualBasic
Imports AmosGraphics
Imports AmosEngineLib. AmosEngine. TMatrixID
Public Class CustomCode
Implements IPlugin
Public Function Mainsub( ) As Integer Implements IPlugin. Mainsub
AddHandler PathDiagrammer. PreFitOptions, AddressOf PreFitOptions
AddHandler PathDiagrammer. PostFitResults, AddressOf PostFitResults
MsgBox( "计算标准化 RMR 的副程序已安装!" & vbCrLf & "Fred Li, 2007",
"Standardized RMR" )
End Function
Private Sub PreFitOptions( ByVal Sem As AmosEngineLib. AmosEngine)
Sem. NeedEstimates( SampleCorrelations)
Sem. NeedEstimates( ImpliedCorrelations)
End Sub
Private Sub PostFitResults( ByVal Sem As AmosEngineLib. AmosEngine,
ByVal ModelName As String, ByVal status As Integer)
Dim N As Integer
Dim i As Integer
Dim j As Integer
Dim DTemp As Double
Dim Sample( , ) As Double
Dim Implied( , ) As Double
Sem. GetEstimates( SampleCorrelations, Sample)
Sem. GetEstimates( ImpliedCorrelations, Implied)
N = UBound( Sample, 1) +1
DTemp = 0
For i = 1 To N - 1
For j = 0 To i - 1
DTemp = DTemp + ( Sample( i, j) - Implied( i, j) ) ^ 2
Next
Next
```

```
DTemp = System. Math. Sqrt( DTemp / (N * (N + 1) / 2))
'Dtemp 代表 SRMR
'讯息窗口中输出结果
Dim message As String
message = "Model: " & ModelName & vbCrLf
If status = 0 Then
message & = "Standardized RMR = " & DTemp. ToString( "#. 0000" )
Else
message & = ( "The model was not successfully . tted. " )
End If
MsgBox( message , ,"      标准化均方根残差" )
End Sub
Public Function Name( ) As String Implements IPlugin. Name
Return "SRMR"
End Function
Public Function Description( ) As String Implements IPlugin. Description
Return "标准化均方根残差"
End Function
End Class
```

本章习题

1. Amos 6.0 所采用的 MCMC 演算法是什么？其基本步骤为何？

2. 在 Bayesian SEM 窗口中，会出现(500 + 70509) * 128 之资讯，这表示 Amos 已进行几次瘦身工作？

第三章 Amos 7.0 新增功能与特色简介

由图 1-1 与图 3-1 比较观之,2006 年问世的 Amos 7.0 之使用者界面与 Amos 6.0 完全相同,但也增加了一些新的贝氏统计与数据填补功能,使其统计分析的功能更加完备。Amos 7.0 新增了七个功能,兹将各新增功能与特色逐一列述如下,以利使用者之运用。

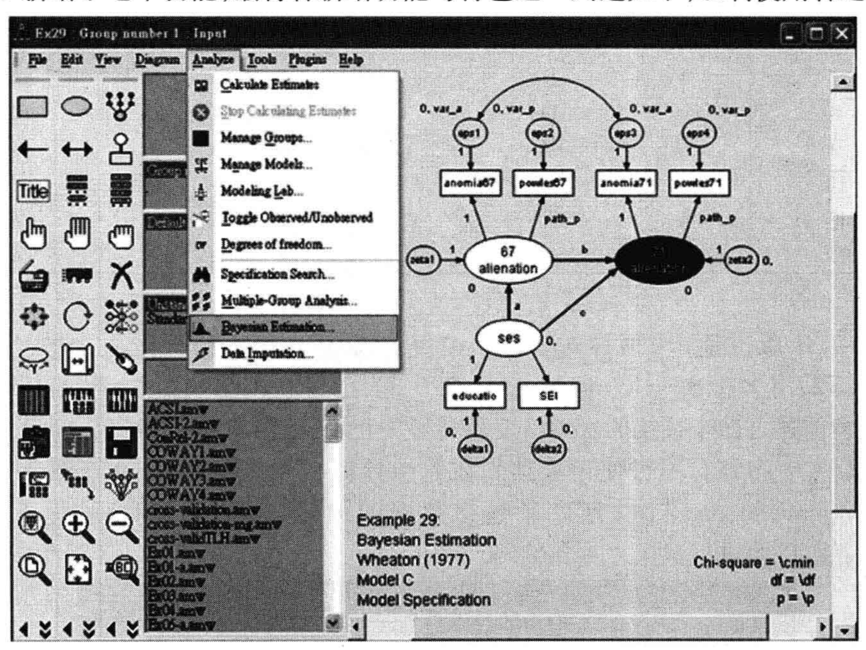


图 3-1 Amos 7.0 使用者界面

第一节 Bayesian SEM 已能处理次序性类别数据与截断数据

Amos 在 7.0 版以前,只能处理等距以上属性之变量,但从 7.0 版开始,Amos 通过 Bayesian SEM 已能处理次序性数据与截断数据。兹将相关之处理步骤,简介如下。

一、次序性类别数据(Ordered-categorical)的数据处理方法

兹以下列态度反应之四点量尺实例,进行数据输入与重新编码的说明:

SD. 非常不同意(Strongly disagree)

D. 不同意(Disagree)

A. 同意(Agree)

SA. 非常同意(Strongly agree)

表 3-1 系前述这四点量尺的次序性类别数据的原始文件(采自 Amos 手册中之 Ex33-a. amw),文件中空白处代表数据缺失。由表之内容知,每一笔原始文件均以字串 SD,D,A,SA 形式键入,此类数据在 SPSS 的数据编辑器中,必须以字串去界定变量属性,才能输入数据。

表 3-1 次序性类别资料的原始输入

	item 1	item 2	item 3	item 4	item 5	item 6
1	A		SA	SD	A	A
2	A		A	SA	SA	SA
3		A	A	A	A	A
4	A	A	A			
5	D	SD			D	
6	SA	SA	A		A	A
7	A	D		A	A	A
8	D	D		SD		SD
9	SA	SA	SA	A		A
10	SA	A	A	SA	SA	
11	A	A	A	A		A
12	SA	SA	A	A		A

Amos 7.0 版以前,均须针对表 3-1 中每一态度反应类别分派一个数值。例如,SD = 1, D = 2, A = 3, SA = 4。但从 Amos 7.0 起,已可利用图 3-2 中[Tools]下之[Data Recode],将前述之四个接连的次序性变量转换为正态性连续性同意量尺,而且量尺的大小亦可由研究者自订。不过使用者须先在原始数据文件连接之时,就于图 3-7 中勾选[Allow non-numeric data],以允许非数字数据的设定,才能使得[Tools]下之[Data Recode]生效,以便进行数据之重新编码。

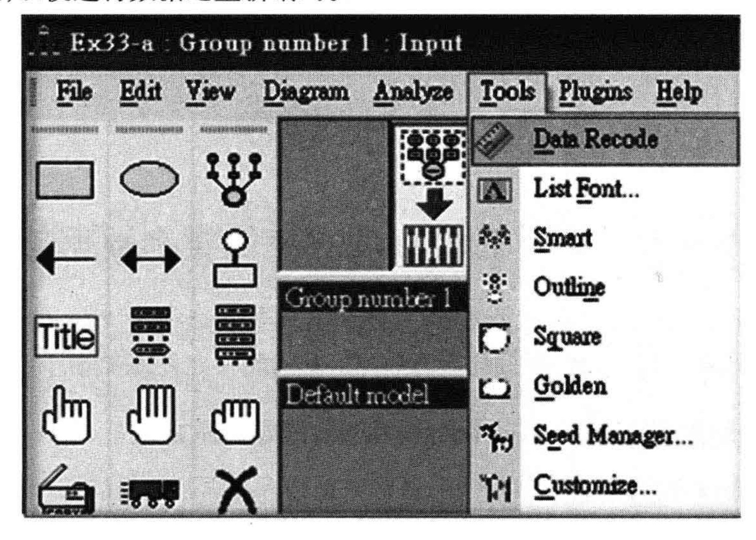


图 3-2 Amos 7.0 数据重新编码表单

进行原始数据的转码，研究者须先点开图 3-2 功能表单[Tools]，再按下[Data Recode]即会出现图 3-3 之数据重新编码的操作窗口，窗口右侧为经转换后之新变量，转换的方法可从窗口中间的选单中点选。数据转码规则有三：①No Recoding：不做转换，②Customized：自订法(只有本法研究者可在[New Value]栏位下直接更改数据，对于的重新编码亦甚便捷)，③Ordered-categorical：视外在的次序性类别指标为连续变量，Amos 会依据您的数据转码方法进行原始数据的转换。图 3-3 窗口之右上角[View Data]，系用来查看各变量之原始数据文件。

假如研究者未针对各变量量尺的边界进行界定，Amos 会自动估计这些相关的边界值(参见图 3-3 底部)，并假定这些值是来自于标准正态分布 N(0, 1)。如欲自行界定边界值，研究者须点选[Ordered-categorical]选目，接着点开[Details]，开启图 3-4 之设定窗口，进行次序性数据的转换。转换时研究者须先使用[Up]与[Down]按钮，进行类别顺序之设定(以符合计分之方向，要先点选欲移动的选目)，再进行与类别相对应的数值定义，进行数值的定义时，请在双箭头右侧的方框中输入选目之边界值。注意，图 3-3 底部显示出，item 1 之量尺类别并未按 SD，D，A，SA 之顺序排列。

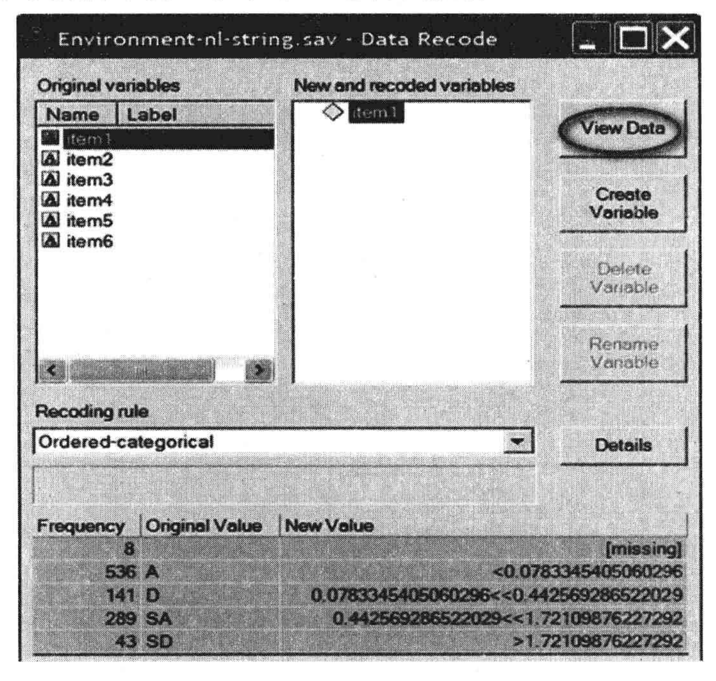


图 3-3 数据转码之操作窗口

另外，图 3-4 中[Ordered-Categorical Details]窗口中，左上角的[Unordered categories]窗口，可供研究者界定缺失值变量。使用时先在[Ordered categories]窗口中，点选缺失值选目(如 SD)再利用 Up 按钮，移上去该窗口中。在本例中 Amos 将[Empty String]视为缺失值。研究者亦可利用[Up]和[Down]键，调整量尺类别之顺序；亦可利用[New Boundary]和[Remove Boundary]增删类别与类别间之边界，其数值的输入须按数值大小依序由小而大从上而下界定之。图 3-4 中[Ordered-Categorical Details]窗口中，左下角的[Ordered-categories]窗口，可供研究者界定边界值。本项新增功能对于反向计分题目的重新编码，亦甚为便利。例如，在图 3-5 中，SD 与 D 被视为无法区分，其潜在分数为 0 以

下，A 选目的潜在分数则介于 0 ~ 1，而 SA 之潜在分数则大于 1 以上。

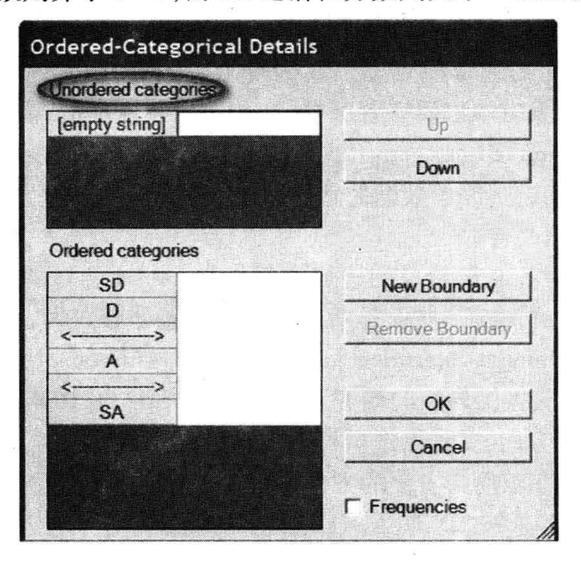


图 3-4　次序性类别资料的细节设定窗口

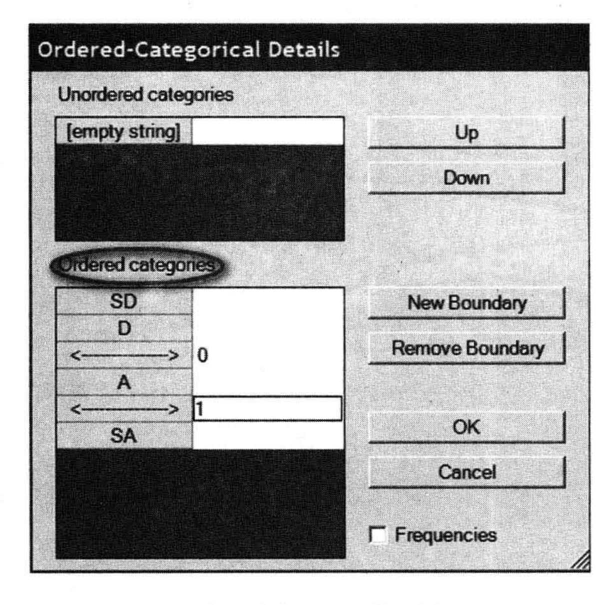


图 3-5　次序性类别数据的边界值设定

经定义过的 item 1 新、旧变量的边界值即可在图 3-6 窗口左下角中显现出来，其余变量（item 2 ~ item 6），以此类推进行边界值之设定。

在图 3-6 中，点选［View Data］即可同时显现新旧变量的数据如表 3-2，表中 0 <；<1 系 A 经转换后之潜在分数，介于 0 与 1 之间；SA 则转换成 >1；D 与 SD 则转换成 <0，即其潜在分数小于 0；而空白缺失值则以 * 表示之。注意此种转换过之潜在分数，只能使用 Bayesian SEM 进行数据分析。此种转换过之类别数据，被视为连续性变量，就可更安心地进行因素分析等统计工作。因为这类转换过之变量被视为连续性变量，相当于边界值之概率等于 0，因此在表 3-2 中并不使用≤ 或≥ 之不等式符号。

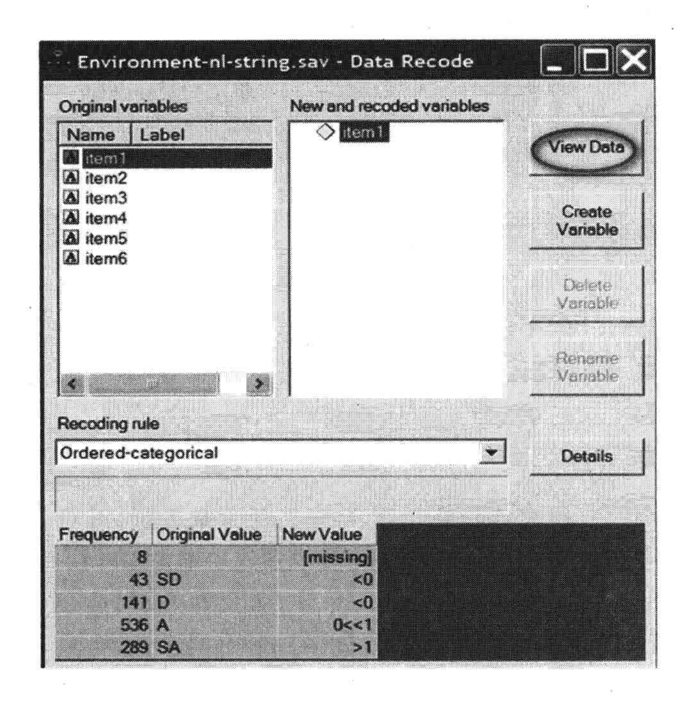


图 3-6　item 1 新、旧变量的设定结果之对照

表 3-2　新旧变量类别与分数的对照表

Data

Original variables

	item1	item2	item3	item4	item5	item6
1	A		SA	SD	A	A
2	A		A	SA	SA	SA
3		A	A	A	A	A
4	A	A	A			
5	D	SD			D	
6	SA	SA	A		A	A
7	A	D		A	A	A
8	D	D		SD		SD
9	SA	SA	SA	A		A
10	SA	A	A	SA	SA	

New and recoded variables

	item1	item2	item3	item4	item5	item6
1	0<;<1	*	>1	<0	0<;<1	0<;<1
2	0<;<1	*	0<;<1	>1	>1	>1
3	*	0<;<1	0<;<1	0<;<1	0<;<1	0<;<1
4	0<;<1	0<;<1	0<;<1	*	*	*
5	<0	<0	*	*	<0	*
6	>1	>1	0<;<1	*	0<;<1	0<;<1
7	0<;<1	<0	*	0<;<1	0<;<1	0<;<1
8	<0	<0	*	<0	*	<0
9	>1	>1	>1	0<;<1	*	0<;<1
10	>1	0<;<1	0<;<1	>1	>1	*

另外，此项新功能将使得 Amos 具有分析是类别变量的依变量的能力，此种分析方法，类似于传统的 Probit 模式分析。当您的依变量是二分的类别变量时，您可以尝试一下 Bayesian SEM 的分析新能力。

二、截断数据的建档、填补与分析方法

截断数据（Censored data）出现于某一变量之测量超过或低于某一阈值，但未知其确切之数值，仅以大于 > 、介于≪或小于 < 某一数据表示之。表 3-3 系七位受试者解决一个问题的时间，其中有两位受试者未在规定的 10 分钟内做完。

表 3-3　截断数据范例

受试者	解决问题的时间
1	6
2	2
3	9
4	>10
5	4
6	9
7	>10

过去于 Amos 6.0 中，研究者须将个案 4 与 7 视为缺失值，或者以一个武断的值如 10，11，12 取代之。任何缺失值的处理方法均会或多或少增大抽样误差，而武断值的填补方法也不是一个好的方法。因此，在 Amos 7.0 中，研究者可依本法充分运用个案 4 与 7 的所有信息，进行统计分析。因含有非数值数据（non-numeric data）如大于 >、介于≪或小于 < 等符号，如以 SPSS 的数据编辑器来编辑资料，须以字串变量设定才能输入资料。此类数据在 Amos 中被视为次序性类别数据，分析此类资料只能以贝氏估计法（传统之 SEM 按钮会失效），而且在 Amos 数据文件设定时，研究者也必须在图 3-7 中勾选[Allow non-numeric data]，才能进行截断数据的相关统计分析。研究者如欲利用传统之 SEM 进行此类数据之分析，可以利用下述贝氏数据填补技术将截断的数据先行填补后，再进行统计分析。

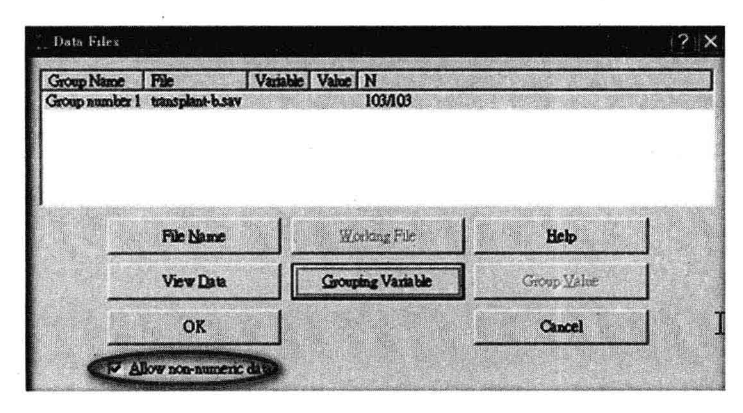


图 3-7　允许非数字数据分析的设定

第二节　提供次序性类别数据与截断数据的缺失值填补法

次序性类别数据与截断数据的测量（过去常视为缺失值）可以提供一个变量更多的资讯。因此，Amos 7.0 可以进行缺失数据、次序性类别数据与截断数据测量之数据填补，乃是一大创举。利用这种填补方式，非数值数据（含截断数据）可以转换为数值，以便无法处理非数据数据的统计方法或软件（如 SPSS）进行数据分析。例如，研究者可以将含

有截断值的数据，利用 Amos 中的贝氏填补法，将截断值从事后预测分布中随机填补后，再利用其他统计方法或软件去分析这些产生出来的数据集，相关之统计分析之步骤，请参阅前面数据填补技术之介绍。

第三节 提供缺失值、次序性类别数据与截断数据的事后预测概率分布

Amos 7.0 可以估计缺失值、次序性类别数据与截断数据等事后预测概率分布(posterior predictive distributions)。在潜在变量的模型中，有三类未知数据：

1. 参数值及其相关之函数(例如，回归及相关系数)。
2. 缺失数据值。
3. 部分缺失数据值，例如，次序性类别数据与截断数据的测量。

在贝氏统计分析中，这三种未知数据的处理方法相同。因为任何一个未知统计量数都可以事后密度函数表示之，以说明这些未知数之可能分布。当资料数据出现缺失或部分缺失时，其事后密度函数称为事后预测概率分布，可作为数据填补之抽样总体。以图 3-8 中 SPSS 数据编辑器中的数据为例，请特别注意 25 和 26 号受试者在 timesqr 变量上的截断数据。

transplant b.sav - SPSS Data Editor

	id	acceptyr	age	priorsur	surgery	time	timesqr
24	24	1969	51.80	0	1	218	14.765
25	25	1969	33.22	0	1	>1799	>42.415
26	26	1969	30.54	0	0	>1400	>37.417
27	27	1969	8.79	0	0	262	16.186
28	28	1969	54.02	0	1	71	8.426
29	29	1969	50.43	0	0	34	5.831
30	30	1969	44.91	0	1	851	29.172

图 3-8 SPSS 数据编辑器中截断数据实例

Amos 之具体操作方法简单说明如下：

配合 Amos 程序档 Ex32-a. amw，在执行 Amos 7.0 [Bayesian estimation] 之后，于 [Bayesian SEM] 的功能表单中，先点选事后预测概率分布的图像(参见图 3-9)，即会出现如图 3-10 左上角之事后预测概率分布的窗口，接着研究者即可针对数据文件中缺失值(在窗口中会以 * 呈现)或截断值(在窗口中会以≫、≪、≪呈现)的数据点，使用鼠标左键按下这些单元格数据点，就可出现图 3-10 右侧之事后概率分布图。

例如，点选图 3-10 中左侧第 26 号受试者在 timesqr 变量上的截断数据(在窗口中会以≫、≪呈现)，即会显示 26 号受试者存活时间的事后预测概率分布与统计结果，如图 3-10 中右侧所示。图 3-10 底部，Amos 提供数个事后预测概率分布制图的选项：(1)多边图(Polygon)；(2)阴影(Shaded)；(3)条图(Histogram)；(4)前 1/3 与后 1/3 分布图(First and last)；(5)追踪(Trace)；(6)自变相关(Autocorrelation)，研究者可善加利用之。

研究者如点选 25 号受试者在 timesqr 变量上的截断数据，并改选[Polygon]即会出现图 3-11。

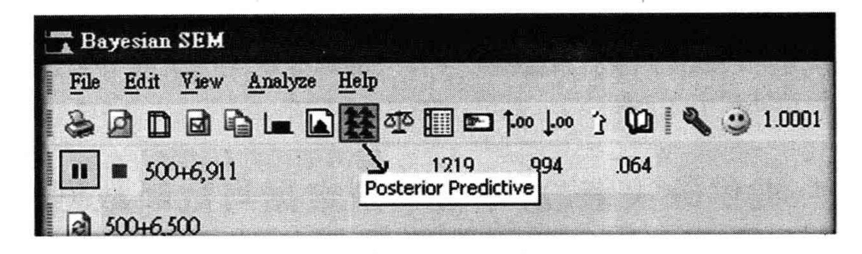


图 3-9 查看事后预测概率分布的按钮

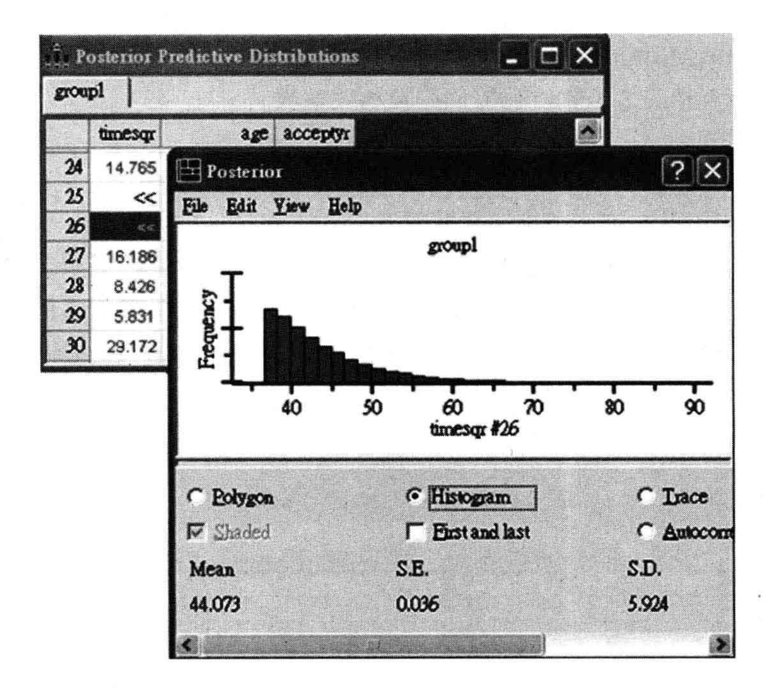


图 3-10 事后预测概率分布的统计结果与分布图

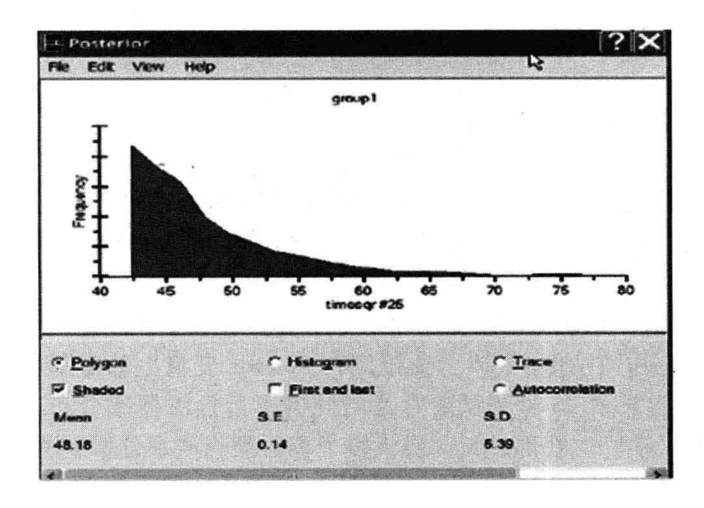


图 3-11 事后预测概率分布图

第四节 增加 Hamiltonian MCMC 演算法

进行此项设定时，请先在 Bayesian SEM 窗口中，点选[View]下的功能表单[Options]，即会出现 Bayesian Options 窗口。Amos 7.0 除了提供 Random Walk Metropolis 演算法(又称醉汉漫步法，因为每一步的大小及方向均为随机)之外，尚提供 Hamiltonian MCMC 演算法(是 Gibbs 与 Metropolis 漫步法之交替运用，试图修正随机漫步之行径)，请参见图 3-12。就数字资料而言，Hamiltonian MCMC 法的演算速度比 Random Walk Metropolis 法快，因为 Hamiltonian MCMC 法会加大在样本空间的移动步伐，促使样本间之相关变小而导致聚敛到目标总体的速度加快。当您的 prior 是[Uniform prior]且您在 Bayesian Options 窗口中点选[Stability]或/与[Admissibility]检验时，建议选择 Random Walk Metropolis 演算法；否则选用 Hamiltonian MCMC 法。另外，在数据填补时则请使用 Random Walk Metropolis 演算法。Hamiltonian MCMC 演算法中的 Step size(epsilon)与 Number of steps(指每一 Hamiltonian MCMC 回圈数中的 Step size)，这两个参数都可利用 MCMC 参数的自动调节[Adapt 按钮]，自动调整其成为合理值，通常研究者不必费心于此。

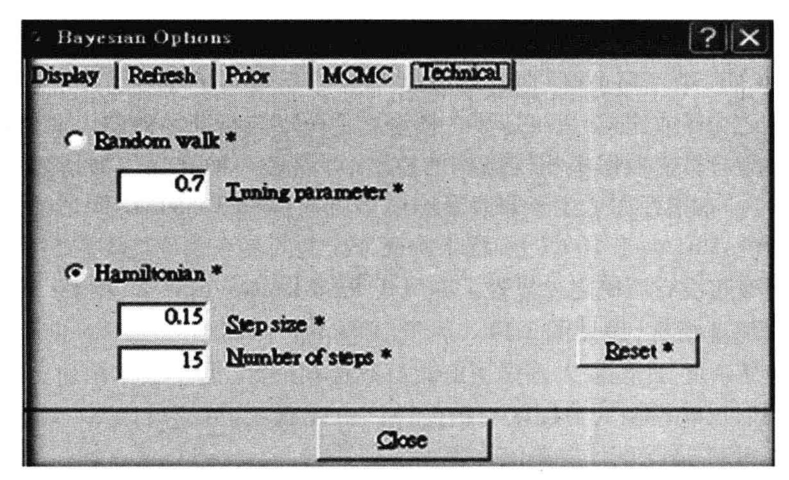


图 3-12 Amos 7.0 MCMC 参数的调节窗口

第五节 MCMC 参数的自动调节

在图 3-13 中 Amos 7.0 提供了一个新的[Adapt button]，可以自动调整 MCMC 演算法的微调参数(使样本之接纳比率介于0.2～0.4)，以提高聚敛之速率。为了改善产生效能，调节钮会分析 MCMC 观察值的历史，进而更新 MCMC 演算法之参数。请谨慎使用它，当您按下此 MCMC 参数的微调钮之后，前面所搜集到的样本会被整个抛弃，再重新抽样。当不须调整 MCMC 演算法的微调参数时，该按钮会变暗而无法再运作。

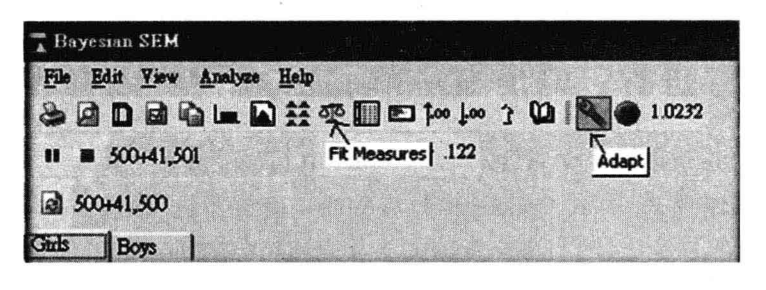


图 3-13 Amos 7.0 MCMC 参数的自动调节与贝氏模型适配度指标按钮

第六节 提供贝氏模型适配度指标

Amos 7.0 可以计算 posterior predictive p 值,它是贝氏导向的 p-value,可用来进行模型的评估。贝氏导向的 p-value,代表着未来数据的卡方值大于或等于观察数据的卡方值的概率。一般来说,当此值小于 0.05 或大于 0.95 时,表示该模型与数据不适配;当此值在 0.5 期望值附近时,表示该模型为适配模型,其合理范围应在 0.05 ~ 0.95(Gelman, Carlin, Stern, & Rubin, 2004; Myung, Karabatsos, & Iverson, 2005)。就数值数据而言,Amos 7.0 亦会计算 DIC(Deviance information criterion)指标,本指标乃是 AIC 与 BIC(这两个指标在 MCMC 中不易求得)之延伸,最适合于事后分布呈现多变量正态分布时,可用来进行模型间之选择或比较,以决定哪一模型之适配度与精简性较佳。最低 DIC 值的模型乃系最能预测与目前数据集相同结构的模型,可说是一种推论力的指标(a measure of generalizability)。本指标亦适用于非嵌套的模型(non-nested models)间之比较,是其一大特色。一般来说,DIC 大于 10 以上,意味着两个模型间具有明确之差异性,如在 10 以下则表示两个模型间没有明确之差异性。在贝氏 SEM toolbar 窗口中,按 Fit Measures 钮即可获得此适配度指标,参见图 3-14。至于[Effective number of parameters]则可用来评估一个模型的复杂度,此值愈大该模型愈能与数据相适配,但却愈不精简。因此,一般来说 DIC 较能同时反映模型的适配度与精简性。由此观之,可知贝氏 SEM 的适配度评鉴似乎比传统之 SEM 适配度评鉴单纯许多。

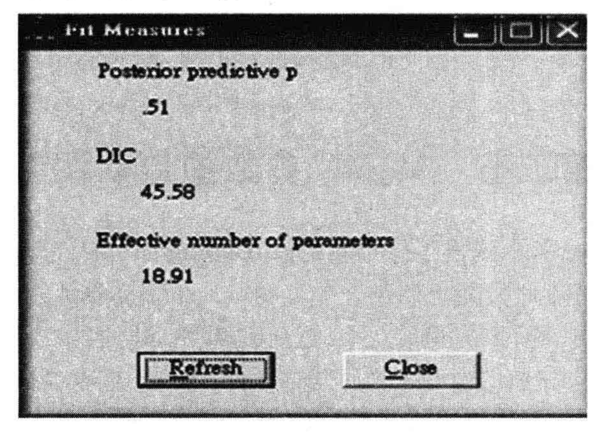


图 3-14 贝氏模型适配度指标显示窗口

第七节　剪贴簿功能的改善

Bayesian SEM 窗口中的图文内容,也可以拷贝及贴至 Microsoft Word 文件中,且仍保留原有表格之格式设定。例如,使用者可以利用[Edit]之下的[Copy]将参数的事后次数分布图(参见图 3-15),贴至 Microsoft Word 文件中。

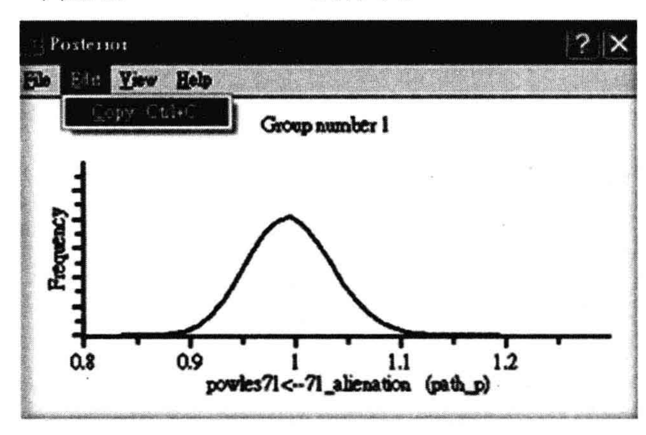


图 3-15　参数的事后次数分布图

本章习题

1. 请至以下网址 http://www.amosdevelopment.com/video/index.htm,观赏有关 Ordered-categorical Data & Censored Data 等等 Amos 教学录影片,并作笔记。

2. 何谓截断数据(Censored data)?请问 Amos 7.0 如何进行截断数据的分析?

3. 请利用 Amos 之[Data Recode]新功能,将表 3-1 的原始文件(采自 Amos 7.0 手册中之 Ex33-a.amw,该文件均以字串 SD、D、A、SA 等形式键入),进行数据之重新编码,重新编码时其边界值依序为 1、2、3。接着再利用此转换过之数据,进行贝氏 SEM 统计分析。请问所得之模式适配度为何?

4. Amos 7.0 所提供的 posterior predictive p 值,研究者如何进行解释与运用?

5. 研究者搜集了过去一年来,喝同一厂牌奶粉的 190 位婴儿之平均每日喝奶量(毒奶粉单位:公升),并检测其有无肾结石(b 表结石者),原始资料如图 3-16 所示。

因本例之依变量系一、二分之类别变量,而预测变量为一连续变量,过去研究者会先利用 SPSS 进行 Probit 分析,输入之相关变量参见图 3-17 中之设定。其中变量“结石 2”为数字型变量,系文字型变量“结石”转换而来;n 系每一组的观察总次数(本例中只有 1 组)。SPSS 仅能分析“结石 2”变量,无法分析文字型变量“结石”,文字型变量“结石”则可以利用 Amos 进行数据分析。

按下[OK]之后,在图 3-18 的 SPSS 的报表中可以看出,毒奶粉食量对结石之回归系数为 3.85492,截距为 -1.40439。由此观之,不喝毒奶粉是不会结石的,而每日喝 1 升的

毒奶粉，必然会结石。请注意，因为 SPSS 的 Probit 模式仅分析整理过之调整数据，如使用原始一对一的个别数据，其跑出来的自由度会不正确，因而导致 χ^2 适配度检验不正确，但其他的参数估计结果仍然正确。

毒奶粉_1.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : 结石2 | 1

	毒奶粉	结石	n	结石2
1	.485	b	1.00	1.00
2	.570	b	1.00	1.00
3	.470	b	1.00	1.00
4	.535	b	1.00	1.00
5	.500	b	1.00	1.00
6	.490	a	1.00	.00
7	.510	b	1.00	1.00
8	.470	b	1.00	1.00
9	.465	a	1.00	.00
10	.500	b	1.00	1.00
11	.545	b	1.00	1.00
12	.435	a	1.00	.00
13	.475	a	1.00	.00
14	.280	b	1.00	1.00
15	.270	b	1.00	1.00
16	.555	b	1.00	1.00
17	.310	b	1.00	1.00
18	.300	b	1.00	1.00
19	.285	b	1.00	1.00
20	.360	b	1.00	1.00

图 3-16　毒奶粉食量与结石者之原始资料

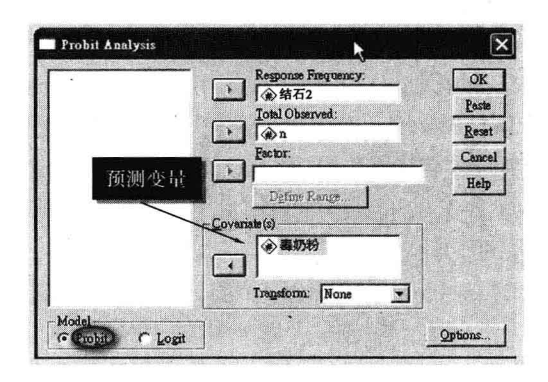
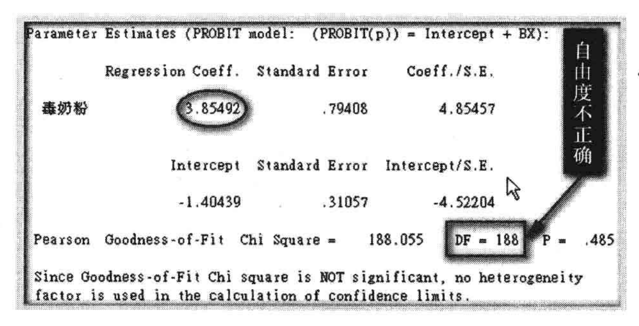


图 3-17　毒奶粉食量与结石者之 Probit 分析之输入窗口

接着，请再利用 Amos 的 Bayesian SEM，进行分析如图 3-16，其相关之路径模型绘制如图 3-19 所示，图中结石变量系一非数字型变量。因二分类别变量在此模型中，会造成模型之不可辨识，必须将残差方差固定为 1。因“结石”系一文字型变量，请在文件连接时，勾选[Allow non-numeric data]，且在分析属性窗口中勾选[Estimate means and intercepts]。得出结果之后，请比较一下 SPSS 与 Amos 之毒奶粉食量对结石之回归系数

大小。两者有重大差异吗？另外，运算出来的模型适配度指标为 0.51（参见图 3-20），代表什么意义？


Parameter Estimates (PROBIT model: (PROBIT(p)) = Intercept + BX):

	Regression Coeff.	Standard Error	Coeff./S.E.
毒奶粉	3.85492	.79408	4.85457

Intercept	Standard Error	Intercept/S.E.
-1.40439	.31057	-4.52204

Pearson Goodness-of-Fit Chi Square = 188.055 DF = 188 P = .485

Since Goodness-of-Fit Chi square is NOT significant, no heterogeneity factor is used in the calculation of confidence limits.

图 3-18　毒奶粉食量与结石者之 Probit 分析结果

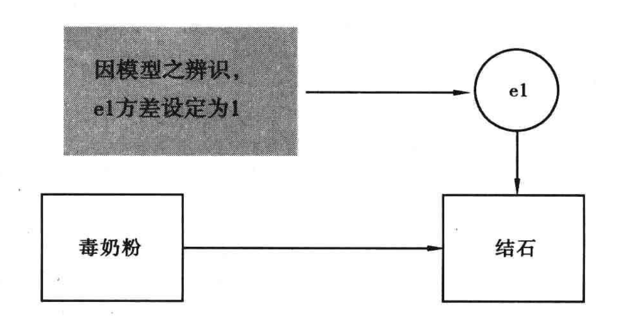


图 3-19　毒奶粉食量与结石者之路径模型

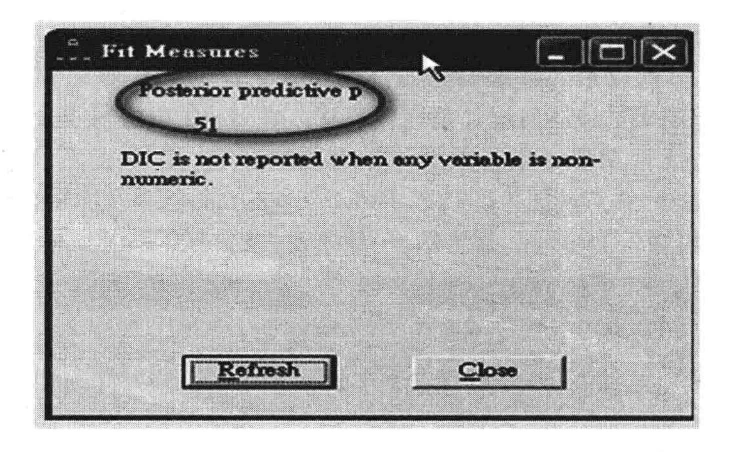


图 3-20　毒奶粉食量与结石者之 Bayesian SEM 的适配度分析结果

第四章 Bootstrapping 的意义及其在模型比较上的运用

本章旨在介绍 Bootstrapping 的意义及其在 Amos 中的操作过程与应用实例分析，以利研究者在学术研究上之运用。

第一节 Bootstrapping 的意义与用途

[Bootstrapping]最常用来产生参数估计值的抽样分布，它是数据自我复制的一种方法，所谓“bootstrap samples”是指以原来的样本为抽样之总体，采用放回随机抽样从原来的样本中抽取同一大小之样本，如此重复抽取所得之样本称为“bootstrap samples”或称为“multiple subsamples of the same size”。接着 Amos 会进行每一 Bootstrap 样本之参数估计，最后再计算每一参数的平均值与标准误。Bootstrapping 不但可以计算传统无法估计的标准误（如 R^2），而且可以不必理会正态性、独立性与大样本之基本假设。因此，Bootstrapping 可以用来估计任何统计量之标准误。在 Amos 中进行 Bootstrapping 的统计分析，须先在分析属性窗口中进行如图 4-1 之参数设定。具体言之，研究者须先点选[Bootstrap]功能表单，并点选图 4-1 中之[Perform bootstrap]选目及所需之适配函数，本例为 Bootstrap ML，再输入 Bootstrap 样本数（一般来说，至少 100 以上），本例为 1000。如输

图 4-1 Bootstrapping 之 Amos 设定窗口

入之数据为相关或协方差矩阵,请加点[Monte Carlo(parametric bootstrap)]。研究者如欲额外输出参数之置信区间,请再点选[Percentile conifdence intervals]。

此外,进行 Bootstrapping 的统计分析时,Amos 不允许数据中有缺失值存在,遇有缺失值时会出现图 4-2 之警告窗口,研究者需要将缺失值之个案加以删除或设法将缺失数据加以填补,才能顺利进行 Bootstrapping 的统计分析。

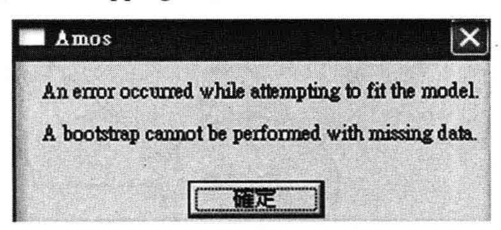


图 4-2　Bootstrapping 时缺失值之警告窗口

第二节　Bootstrapping 统计分析之实例示范

兹以吴丽华(2007)的硕士论文资料为例，说明如何利用 Amos 之[Bootstrapping]功能,来估计理论模型中各统计量之标准误,其理论模型参见图 4-4。该笔资料因具有部分缺失值,无法直接进行 Bootstrapping 的统计分析，研究者可使用 Amos 或 SPSS 的数据填补法，以进行缺失数据的填补。以下之 Bootstrapping 统计分析,系根据 SPSS 的回归填补法所建立数据,操作时请依下列步骤:Analyze ⟶Missing Value Analysis ⟶Estimation 之设定⟶Regression 之设定。研究者如欲将填补过之文件存为不同文件,请在图 4-3 中点选[Save Completed Data],并按下 File 按钮,在开启之窗口中输入新文件名,以便存档。

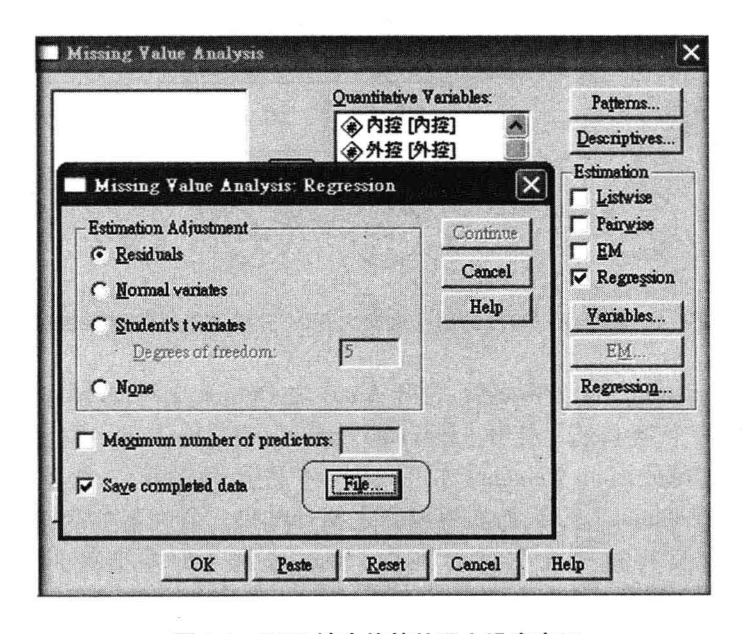


图 4-3　SPSS 缺失值的处理之设定窗口

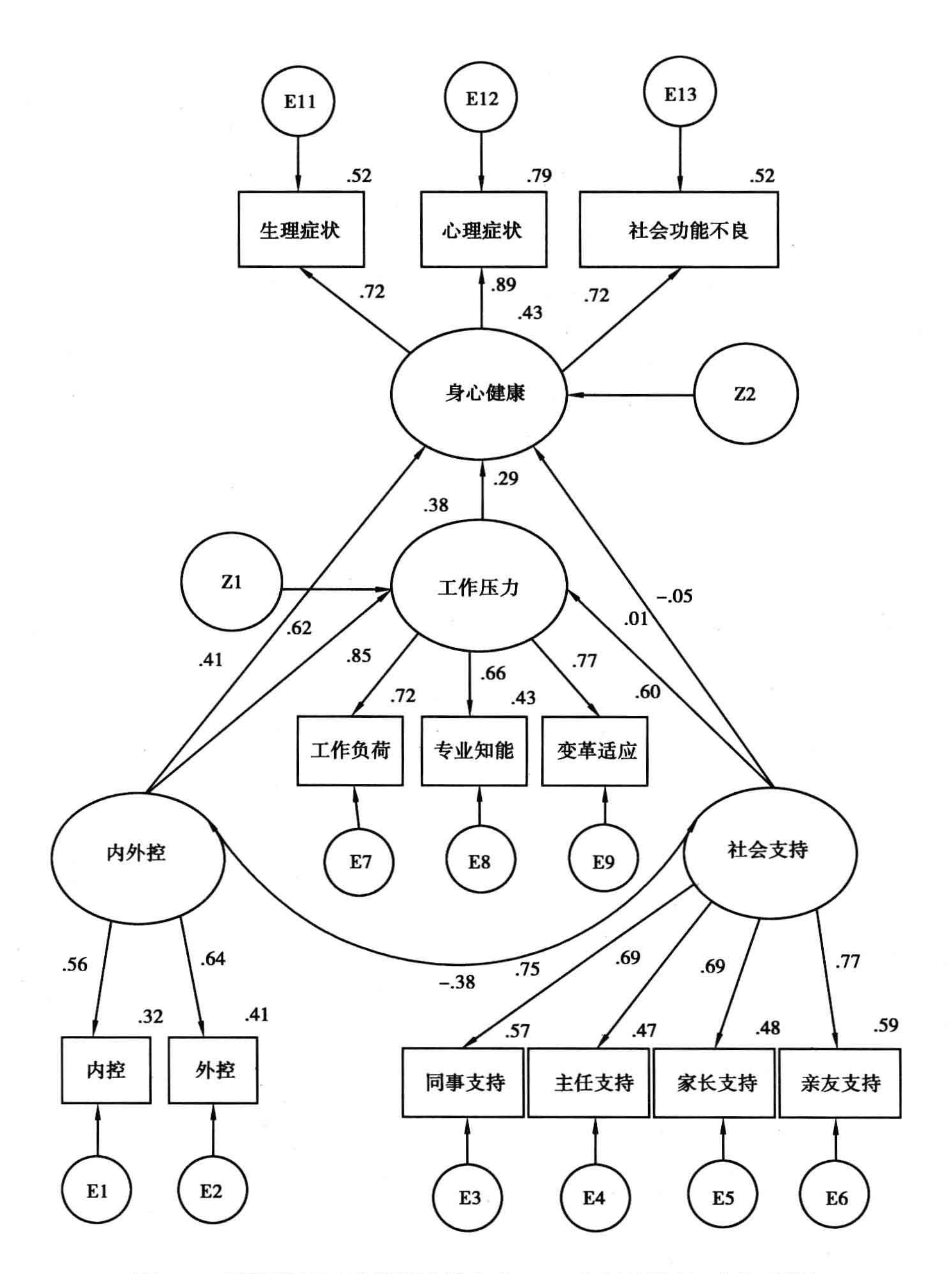


图 4-4　吴丽华(2007)的硕士论文之 SEM 分析结果(标准化系数)

研究者于 Bootstrapping 分析之后,打开 Amos 之 Output 文件,如图 4-5 之内容。接着在左上角的第一个窗口内,点开[Estimate]选目,在[Scalars]中,点选[Regression Weights]。这些[Regression Weights]亦可在图 4-5 左上角的第二个窗口内(Estimates/Bootstrap),点选[Estimate]选目,而获得相同之统计结果。研究者如欲获得 Bootstrapping 分析之回归系数及其相关之估计标准误,请点选图 4-5 左上角的第二个窗口内(Estimates/Bootstrap)的[Bootstrap Standard Errors]选目,就会出现如图 4-6 之 Bootstrapping 分析之回归系数及其相关之统计量。图 4-6 中 Bias 表示两种统计方法之差异量,亦即图 4-5 中传统的各参数的 Estimate 与图 4-6 中 Bootstrapping 各参数的 Mean 相

对应的差异值。例如,第一对参数(工作压力与社会支持的回归系数)的 Bias 等于 -0.001(0.004 - 0.005)。

注意图 4-7 中之 R^2 系数估计值,过去传统的 SEM 分析无法计算这些参数估计值的标准误。

Regression Weights: (Group number 1 - Default model)

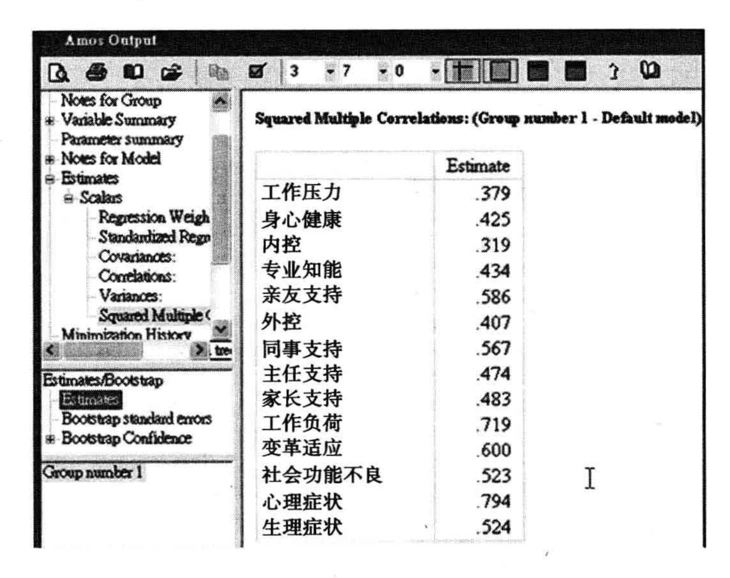
			Estimate	S.E.	C.R.	P	Label
工作压力	<---	社会支持	.005	.031	.149	.882	
工作压力	<---	内外控	.328	.036	9.216	***	
身心健康	<---	工作压力	.251	.060	4.183	***	
身心健康	<---	社会支持	-.021	.023	-.923	.356	
身心健康	<---	内外控	.185	.039	4.716	***	
心理症状	<---	身心健康	1.000				
变革适应	<---	工作压力	1.000				
工作负荷	<---	工作压力	1.000				
家长支持	<---	社会支持	.455	.025	18.143	***	
主任支持	<---	社会支持	.482	.027	17.921	***	
同事支持	<---	社会支持	.414	.021	20.103	***	
亲友支持	<---	社会支持	.460	.022	20.521	***	
专业知能	<---	工作压力	.840	.048	17.465	***	
社会功能不良	<---	身心健康	.555	.030	18.493	***	
生理症状	<---	身心健康	.967	.052	18.512	***	
内控	<---	内外控	.317	.027	11.912	***	
外控	<---	内外控	.360	.028	12.941	***	

图 4-5 传统 SEM 分析的回归系数及其相关的统计量

Regression Weights: (Group number 1 - Default model)

Parameter			SE	SE-SE	Mean	Bias	SE-Bias
工作压力	<---	社会支持	.036	.001	.004	-.001	.001
工作压力	<---	内外控	.042	.001	.328	.000	.001
身心健康	<---	工作压力	.070	.002	.245	-.007	.002
身心健康	<---	社会支持	.027	.001	-.018	.003	.001
身心健康	<---	内外控	.047	.001	.191	.005	.001
心理症状	<---	身心健康	.000	.000	1.000	.000	.000
变革适应	<---	工作压力	.000	.000	1.000	.000	.000
工作负荷	<---	工作压力	.000	.000	1.000	.000	.000
家长支持	<---	社会支持	.027	.001	.454	-.001	.001
主任支持	<---	社会支持	.031	.001	.481	-.001	.001
同事支持	<---	社会支持	.025	.001	.414	.000	.001
亲友支持	<---	社会支持	.027	.001	.459	-.001	.001
专业知能	<---	工作压力	.053	.001	.840	.000	.002
社会功能不良	<---	身心健康	.039	.001	.554	-.001	.001
生理症状	<---	身心健康	.052	.001	.970	.003	.002
内控	<---	内外控	.035	.001	.317	.000	.001
外控	<---	内外控	.034	.001	.359	-.001	.001

图 4-6 Bootstrapping 分析的回归系数及其相关的统计量



Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
工作压力	.379
身心健康	.425
内控	.319
专业知能	.434
亲友支持	.586
外控	.407
同事支持	.567
主任支持	.474
家长支持	.483
工作负荷	.719
变革适应	.600
社会功能不良	.523
心理症状	.794
生理症状	.524

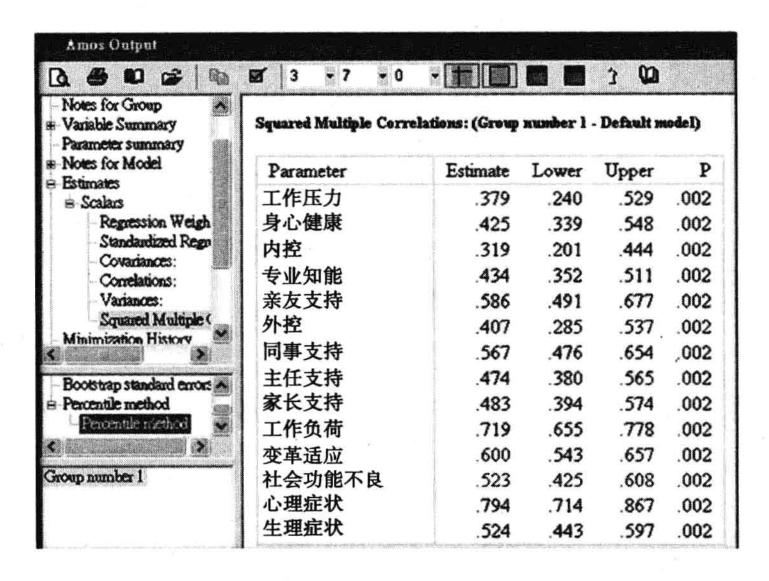
图 4-7　传统 SEM 分析的 R^2 系数估计值

注意图 4-8 中这些 Bootstrapping 分析的 R^2 系数估计值，已可计算相关参数估计值的标准误（SE）。另外，研究者如欲额外输出参数之置信区间，可再点选图 4-1 中的 [Percentile confidence intervals]（参见图 4-9）。图 4-9 中 Bias 的意义如同前述，系指传统 R^2 系数估计值与 Bootstrap R^2 系数估计值间的差异。

Squared Multiple Correlations: (Group number 1 - Default model)

Parameter	SE	SE-SE	Mean	Bias	SE-Bias
工作压力	.072	.002	.385	.006	.002
身心健康	.054	.001	.435	.010	.002
内控	.063	.001	.322	.002	.002
专业知能	.042	.001	.435	.001	.001
亲友支持	.047	.001	.586	.000	.001
外控	.068	.002	.407	.000	.002
同事支持	.047	.001	.567	.000	.001
主任支持	.047	.001	.473	-.001	.001
家长支持	.045	.001	.484	.001	.001
工作负荷	.032	.001	.719	.001	.001
变革适应	.030	.001	.600	.000	.001
社会功能不良	.046	.001	.522	-.001	.001
心理症状	.039	.001	.795	.000	.001
生理症状	.038	.001	.526	.002	.001

图 4-8　Bootstrapping 分析的 R^2 系数估计值与标准误



Parameter	Estimate	Lower	Upper	P
工作压力	.379	.240	.529	.002
身心健康	.425	.339	.548	.002
内控	.319	.201	.444	.002
专业知能	.434	.352	.511	.002
亲友支持	.586	.491	.677	.002
外控	.407	.285	.537	.002
同事支持	.567	.476	.654	.002
主任支持	.474	.380	.565	.002
家长支持	.483	.394	.574	.002
工作负荷	.719	.655	.778	.002
变革适应	.600	.543	.657	.002
社会功能不良	.523	.425	.608	.002
心理症状	.794	.714	.867	.002
生理症状	.524	.443	.597	.002

图 4-9　R^2 置信区间(.95)

第三节　Bootstrapping 在模型比较上的运用

根据 Linhart 与 Zucchini(1986)的模型选择方法,Amos 使用下列四个步骤进行模型的选择(引自《Amos 4.0 操作手册》,pp. 369-370):

1. 采用放回随机抽样,从原来的样本(视为 bootstrap sampling 的总体)中抽取同一大小之 bootstrap 样本;
2. 在每一 bootstrap 样本上,进行各竞争模型的适配性分析,并计算 bootstrap 样本与 bootstrap 总体间的适配函数(implied sample moment vs population moment)之差异值;
3. 计算前一步骤的适配函数差异值的平均数;
4. 比较各模型的平均差异值,出现最小的平均适配函数差异值,即为最佳模型。

以下将以林姿饴(2005)的三个"数常识"理论模型(五因素、四因素及四因素修正模型)为例,逐步简介 Amos Bootstrapping 在模型比较上的应用。首先,打开 Amos 的 Bootstrapping 设定窗口,设定相关的统计方法,参见图 4-10 的设定,请特别注意打勾处的设定。

如非原始数据文件,请同时点选[Monte Carlo(parametric bootstrap)],否则会出现图 4-11的警告窗口。

为了让三个理论模型均使用相同的抽样样本集以利比较,研究者必须在图 4-12 的窗口内,设定相同的随机数起始值。

第一个数常识的模型为五因素模型,如图 4-13 所示。

表 4-1 为模型一之 ML 极小化差异函数的分布图(抽样样本隐含协方差矩阵 vs 抽样总体协方差矩阵)。

表 4-1 系这 1000 个 bootstraps 的次数分布图,其平均 χ^2 的值为 713.957,本统计量反

Analysis Properties
Estimation | Numerical | Bias | Output | Bootstrap | Permutations | Random # | Title
Perform bootstrap
1000 Number of bootstrap samples
Percentile confidence intervals
95 PC confidence level
Bias-corrected confidence intervals
90 BC confidence level
Bootstrap ADF
Monte Carlo (parametric bootstrap)
Bootstrap ML
Report details of each bootstrap sample

图 4-10 Bootstrapping 设定窗口

Amos
An error occurred while attempting to fit the model.
In the "Analysis Properties" window, if "Perform bootstrap" is checked and "Monte Carlo (parametric bootstrap)" is unchecked, the data file must contain raw data.
Similarly, if you use the Bootstrap method without using the MonteCarlo method, the data file must contain raw data.
确定

图 4-11 非原始数据文件警告窗口

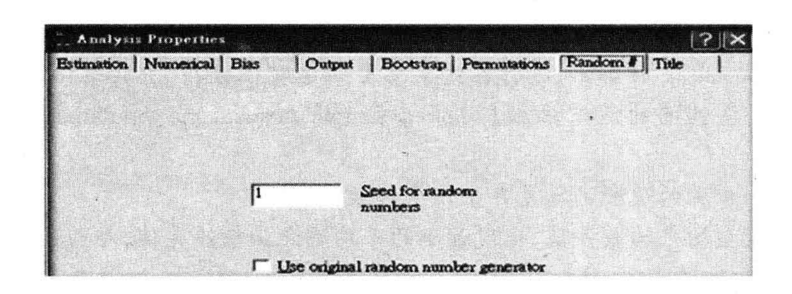


图 4-12 随机数起始值的设定

映出隐含样本矩(implied sample moment,也译样本动差)与总体矩(population moment,也译总体动差)间的极小化差异量。

第二个数常识的模型为四因素模型,如图 4-14 所示。

表 4-2 为模型二的 ML 极小化差异函数的分布图(抽样样本隐含协方差矩阵 vs 抽样总体协方差矩阵)。

第三个数常识的模型为四因素修正模型(Basic 程序设计,请参见附录二),如图 4-15 所示,部分误差项具有相关:e9 vs e55 和 e14 vs e21。

表 4-3 为模型三的 ML 极小化差异函数的分布图(抽样样本隐含协方差矩阵 vs 抽样总体协方差矩阵)。

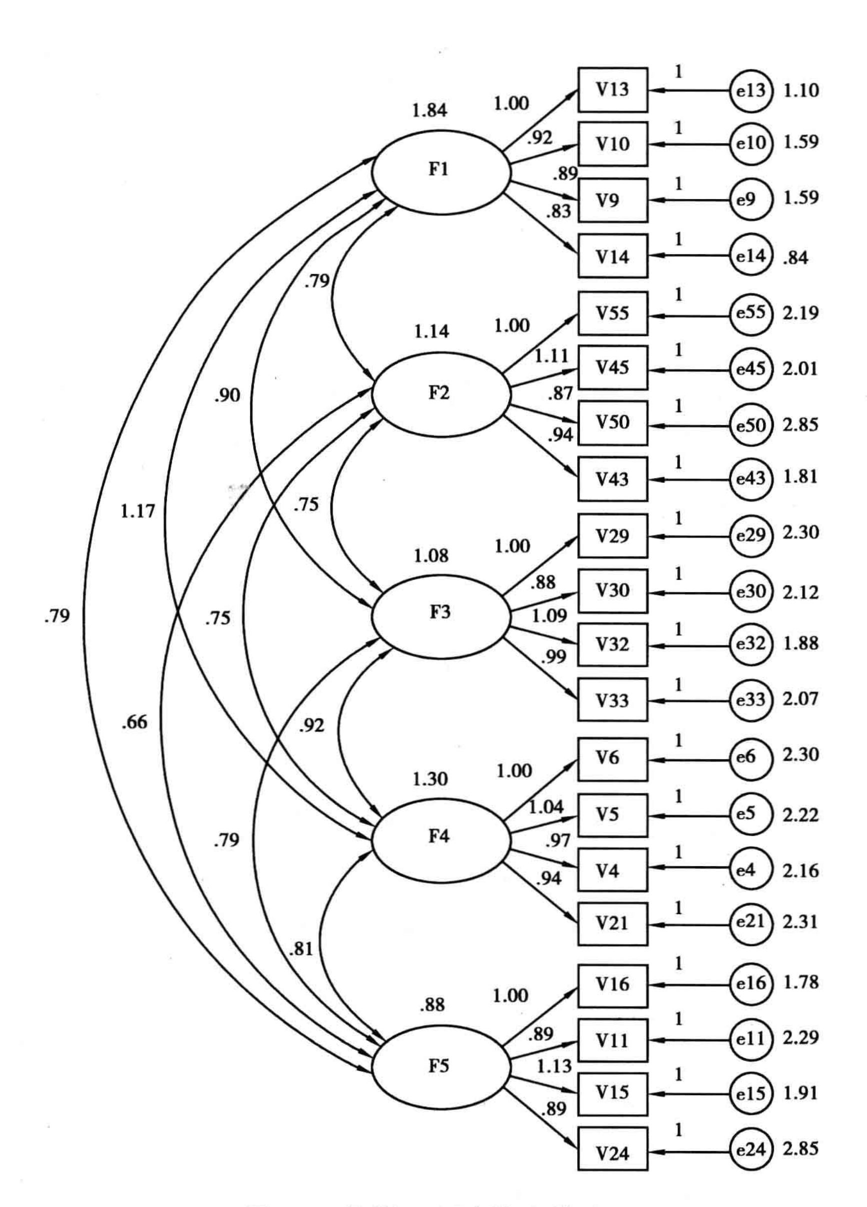


图 4-13 数常识五因素模型:模型一

表 4-1　模型一之 ML 极小化差异函数的分布图

```
                          |-------------------------------
                  680.368 |*
                  687.840 |**
                  695.312 |********
                  702.784 |**************
                  710.257 |********************
                  717.729 |**************
N = 1000          725.201 |***********
Mean = 713.957    732.673 |*******
S. e. = .416      740.145 |***
                  747.617 |*
                  755.090 |*
                  762.562 |*
                  770.034 |
                  777.506 |
                  784.978 |*
                          |-------------------------------
```

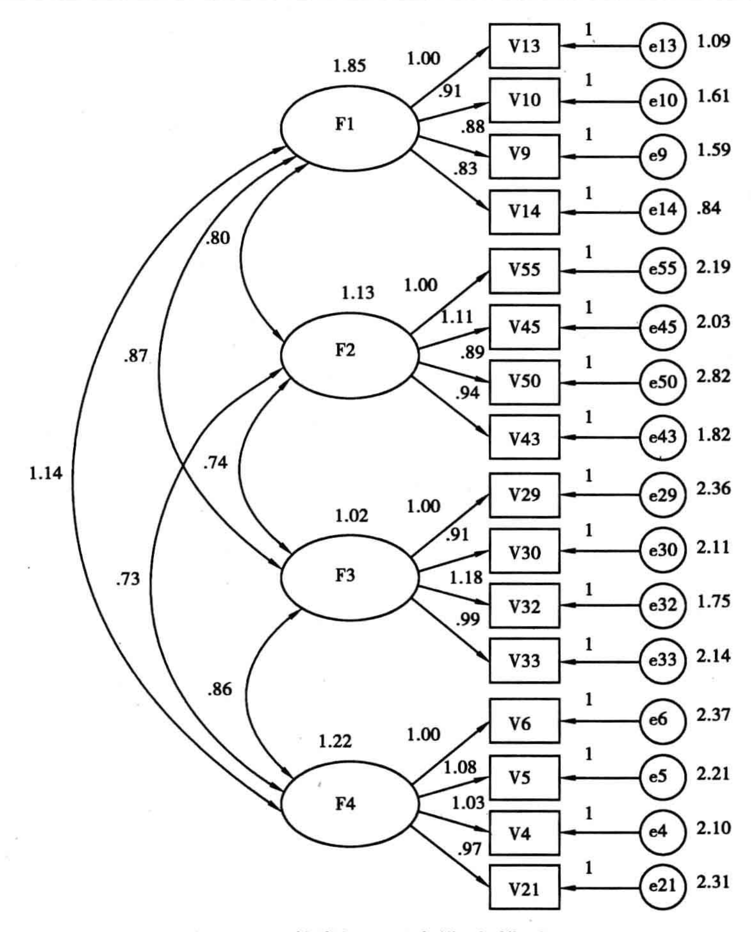


图 4-14　数常识四因素模型:模型二

表 4-2 模型二之 ML 极小化差异函数的分布图

		|----------------------------
	469.905	| *
	474.823	| *
	479.741	| * * *
	484.658	| * * * * * * *
	489.576	| * * * * * * * * * * * * * * *
	494.494	| * * * * * * * * * * * * * * * * * * *
N = 1000	499.412	| * * * * * * * * * * * * * * * *
Mean = 500.549	504.330	| * * * * * * * * * * * * * * * *
S. e. = .356	509.248	| * * * * * * * * * * *
	514.166	| * * * * * * *
	519.084	| * * * * *
	524.002	| * * *
	528.920	| * *
	533.838	| *
	538.756	| *
		|----------------------------

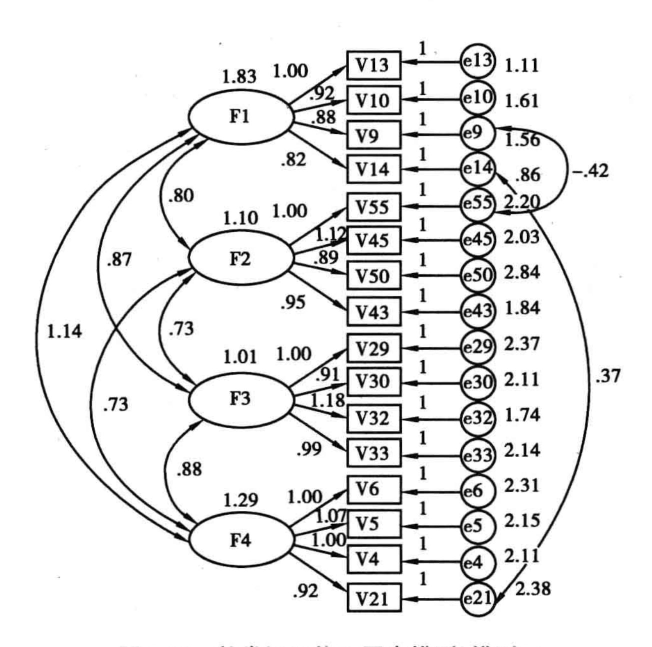


图 4-15 数常识四修正因素模型:模型三

表 4-3 模弄三之 ML 极小化差异函数的分布图

```
                                      |----------------------------
                         416.630      | *
                         421.629      | *
                         426.627      | * *
                         431.625      | * * * * * *
                         436.623      | * * * * * * * * * * * *
                         441.621      | * * * * * * * * * * * * * * * * * * *
N = 1000                 446.620      | * * * * * * * * * * * * * * * * * *
Mean = 449.155           451.618      | * * * * * * * * * * * * * *
S. e. = .363             456.616      | * * * * * * * * * * * * *
                         461.614      | * * * * * * * * *
                         466.613      | * * * * *
                         471.611      | * * *
                         476.609      | * *
                         481.607      | *
                         486.605      | *
                                      |----------------------------
```

兹将前述三个数常识之理论模型的主要适配度指标及各模型的 χ^2 平均差异值，整理如表 4-4 以便进行竞争模型的比较：

表 4-4 三个数常识之竞争模型的比较：Bootstrapping 法

模型	χ^2平均值	AIC	BCC	BIC	CAIC
1	713.957	741.559	745.07	963.13	1013.13
2	500.549	521.12	523.26	689.51	727.51
3	449.155	471.58	473.8	4648.84	688.84

根据前述 Linhart 与 Zucchini(1986)的模型选择方法，利用 Amos 的 Bootstrap 内建功能，分析结果如表 4-4 所示：就 χ^2 与适配度指标来看，很显然第三个理论模型为最佳的理论模型。不过，因为 χ^2 与适配度指标常无法反应模型的正确性与参数的合理性，研究者仍须注意该模型的理论依据与相关参数估计值的合理性。

本章习题

1. 何谓 Bootstrapping?
2. 利用 Bootstrapping 可估计任何统计量之标准误吗?

第五章　潜在特质变化分析

探讨研究变量的稳定性与变迁,是社会科学的研究核心之一。例如,教育学者欲探究学生学习成长的问题(Kim, Murry, & Brody, 2001; Shay, & Gomez,2002),心理学者想了解儿童的认知发展情形(Coscia, Ris, & Succop, 2003;Greene, Way, 2005),社会学者想了解犯罪率的变化问题(Hipp, Bauer, Curran,& Bollen, 2004)。利用传统的单变量或多变量重复量数方差分析,进行纵贯性数据分析,通常偏向整体性的成长分析,而且受到一些统计基本假设上的限制。近十年来在线性结构方程模型中,出现了新的分析方法:潜在特质变化分析(latent change analysis,简称 LCA),而且似有逐渐受重视的趋势。此种潜在特质变化分析,亦常见于分析饮酒、吸烟、吸毒、偏差行为、反社会行为,以及家庭咨询的研究上(Windle, 2000)。

第一节　潜在特质变化分析的意义与重要性

潜在特质变化分析,常用来探究个体内潜在特质变化的情形(Intraindividual differences in change)或这些变化的情形在个体间的差异性,有时也被称为成长曲线模型(Growth curve modeling)或潜在特质成长分析(Latent growth models)。LCA 本质上是SEM 协方差—平均数结构分析或多层次分析(Multilevel models)的特例,旨在利用同一受测者接受多次的测量,分析个体内潜在特质变化平均变化的大小与类型,它可以避免传统重复量数之球形假设(Assumption of sphericity)与测量无误差的基本假设。一般协方差—平均数结构分析的结构回归模型为:

$$\eta = \alpha + B\eta + \Gamma\xi + \zeta$$

而其测量模型为:

$$y = \tau_y + \Lambda_y\eta + \varepsilon$$

$$x = \tau_x + \Lambda_x\xi + \delta$$

上列公式中的 τ_x,τ_y 截距向量,是反映自变量与依变量之测量水平的差异的效果,而 α 截距向量是反映潜在变量平均数的大小,但在传统的共变结构 SEM 分析中会将截距设定为0($\tau_x = 0$,$\tau_y = 0$,$\alpha = 0$)。因为 LCA 是 SEM 之特例,以图 5-5 单因子潜在特质变化分析为例(截距为0),则只需用到测量模型:

$$y = \tau_y + \Lambda_y\eta + \varepsilon$$

将上式化成矩阵方式呈现：$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ L2 \\ L3 \\ L4 \end{bmatrix} [\text{Time}] + \begin{bmatrix} e1 \\ e2 \\ e3 \\ e4 \end{bmatrix}$，式中 Time 系时间因子，主宰着个体本身潜在特质变化的趋势。

另外，为何潜在特质的变化分析相当重要？以学校教育为例，学生进入学校接受教育，主要目的在于学习，在于成长，在于发展，在于改变，而这些个体内及个体间学习前后的成长状态与影响因素，正是学校效能或品质的核心指标(Willett，1989)。

第二节　Amos 的 LCA 分析路径图的操作

Amos 为了便利 LCA 分析者制作路径图，特别提供[Growth Curve Model]增益集，其操作步骤简述如下：

一、打开 Amos Plugins 表单(如图 5-1)，点选[Growth Curve Model]

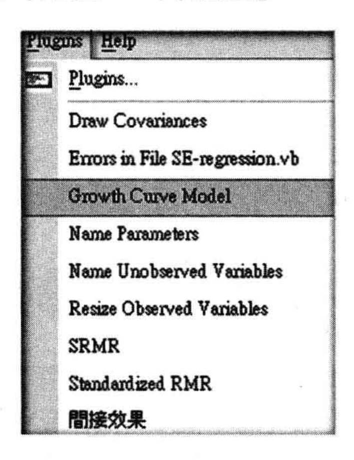


图 5-1　Amos 之 Plugins 表单

二、在图 5-2[Growth Curve Modeling]窗口中，输入重复测量的数目后，按下 OK

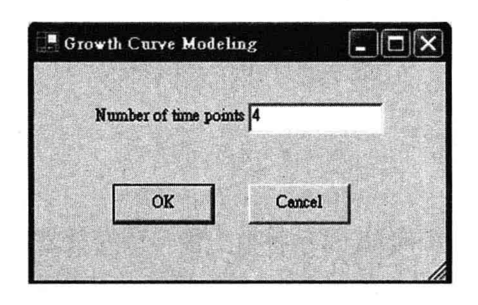


图 5-2　Amos 的 Growth Curve Modeling 窗口

三、在出现的图 5-3 预设路径图上，进行指标变量名称、路径系数、潜在变量名称等的输入与修改

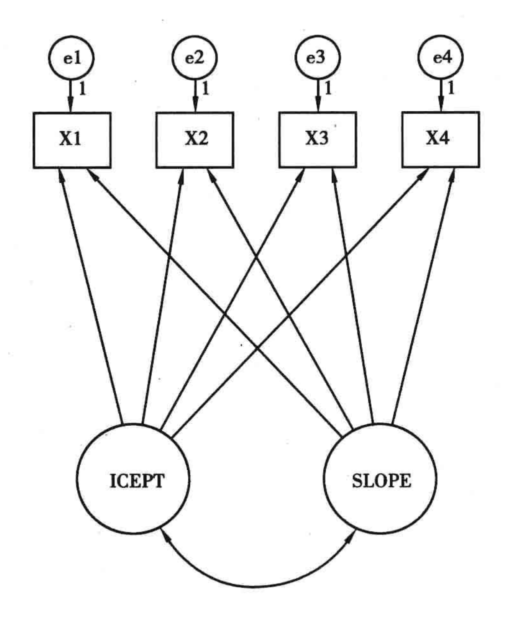


图 5-3 Amos 中预设 Growth Curve Modeling 路径图

四、在图 5-4 的分析属性窗口中，点选[Estimate means and intercepts]

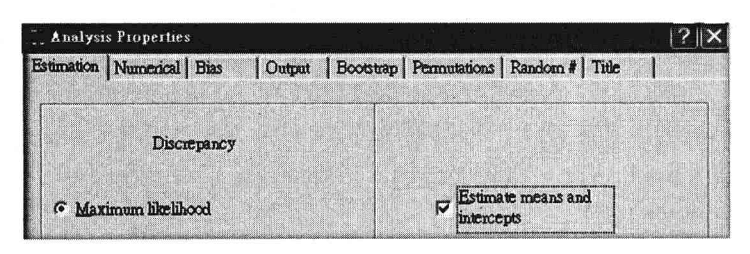


图 5-4 分析属性窗口

五、连接数据文件

请在功能表单[File]下，利用[Data File]窗口设定之。

六、执行[Calculate Estimates]，进行数据分析

请在功能表单[Analyze]之下，执行参数估计(参见图 1-2)。

第三节 LCA 在统计及研究设计上的要求

为了让协方差—平均数结构分析中的参数能被独一无二地估计出来，LCA 除了必须遵守协方差结构分析中的六大规定外，尚需符合方差、协方差与平均数的五大运算原则

(Raykov & Marcoulides, 2006)及 LCA 的基本要求。

一、协方差结构分析中参数性质的六大基本规定

1. 所有自变量的方差为模型参数；
2. 所有自变量间的协方差为模型参数；
3. 所有连接潜在变量与相关指标的因素负荷量为模型参数；
4. 所有连接预测变量与效标的回归系数为模型参数；
5. 所有依变量间(或自变量与依变量间)的方差与协方差均为模型参数；
6. 模型中每一潜在变量的测量单位须加以设定。

上述这些模型参数,除非由于研究或模型之需要加以固定或限制之外,均为等待估计的自由模型参数。另外,由于残差项及测量误差项被视为完全随机的变量,因此他们的期望值均等于 0。

二、方差、协方差与平均数的五个运算原则

1. Cov(X, X) = Var(X)；

2. Cov(aX + bY, cZ + dU) = acCov(X, Z) + adCov(X, U) + bc(Y, Z) + bdCov(Y, U)；

3. Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y),假如 X, Y 具有相关；

4. Var(aX + bY) = a^2Var(X) + b^2Var(Y) ,假如 X, Y 相关等于 0；

5. Mean(aX ± bY) = aMean(X) ± bMean(Y)。

根据这五大运算原则,即可利用模型中的参数,计算出变量的平均数、方差与协方差。兹以图 5-5 之单因子 LCA 模型为例,说明 Cov(Y1, Y2) 与 Var(Y1)、Var(Y2) 与模型参数的关系。利用运算原则 1 ~ 4, Cov(Y1, Y2) = Cov (L1 · Eta + e1, L2 · Eta + e2) = L1L2Var (Eta), Var (Y1) = $(L1)^2$Var (Eta) + Var (e1), Var (Y2) = $(L2)^2$Var(Eta) + Var(e2)。

其余的关系,以此类推,不再赘述。根据前述协方差结构分析中参数性质的基本规定 6,将 Y1 与 Eta 的路径系数固定为 1,而使得潜在变量的平均数等于 Y1 的平均数,亦即 Mean(Eta) = Mean(Y1)。换言之,我们将以 Y1 作为基准线,以评估后续重复量数随着时间之改变比率。

又如,以 Y2 = L2η + e2 为例,利用方差、协方差与平均数的运算原则 5, Y2 的平均数可由下式求得：Mean(Y2) = L2Mean(Eta) + Mean(e2),但因 Mean(e2) = 0,所以 M(Y2) = L2Mean(Eta)。换言之, Y2 的平均数可由潜在变量的平均数与因素负荷量(L2) 的交乘积求得。但因 Mean(Eta) = Mean(Y1), M(Y2) = L2Mean(Eta) 的关系等于 Mean(Y2) = L2Mean(Y1),将 Mean(Y1) 移项可得。其余 $L2 = \frac{Mean(Y2)}{Mean(Y1)}$、$L3 = \frac{Mean(Y3)}{Mean(Y1)}$、$L4 = \frac{Mean(Y4)}{Mean(Y1)}$,亦即这些路径系数为两平均数之比值,而又均以第一次测量 Y1 之平均数做分母。由此可知,图 5-5 单因子 LCA 模型的精心规划,使得因素负荷量,能够反映出潜在特质随着时间之改变比率。

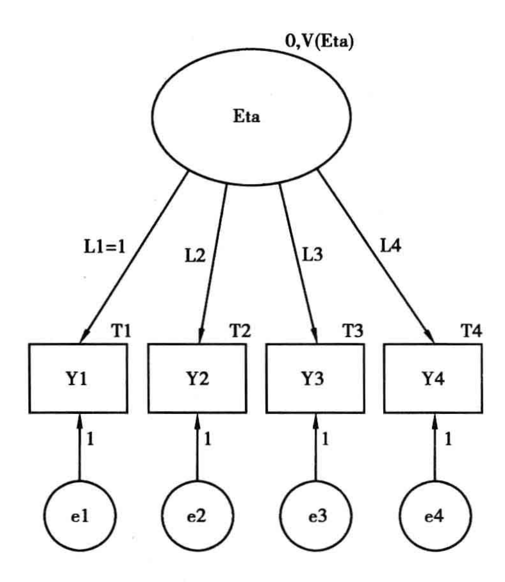


图 5-5　单因子 LCA 模型

三、LCA 之基本要求

在研究设计上，潜在特质成长曲线分析须符合以下几大要求（Kline，2004）：

1. 至少要在不同时间测量三次以上；
2. 测量的变量至少要有等距测量的属性；
3. 每次的测量须全在测同一特质，且未标准化；
4. 所有受试者在各个搜集时间点上，均须接受测量。

以下将针对单、双因子潜在特质变化分析、双因子潜在特质变化之共变分析与多重指标潜在特质成长模型，逐一进行简单说明。

第四节　单因子潜在特质变化分析

单因子潜在特质变化分析只涉及测量水平因子（time factor）的分析，因为此模型将时间因素视为主宰个体内潜在特质变化的主因，因而常被称为成长曲线模型（growth curve model）。相关之路径设计如图 5-6 所示，模型中第一次测量的因素负荷量设定为 1，乃是研究者希望将第一次测量当作基准点，俾便与后续测量做比较，其他的路径系数（L2，L3，L4）则开放估计。

由图 5-6 知，此单因子 LCA 线性模型的测量回归模型可罗列如下：

$$Y1 = 1.0 * Time + e1;$$
$$Y2 = L2 * Time + e2;$$
$$Y3 = L3 * Time + e3;$$
$$Y4 = L4 * Time + e4。$$

利用前述之方差、协方差与平均数的运算原则 5，可推得：Mean（Y1） = 1.0 * Mean（Time） + Mean（e1），因为 Mean（e1） = 0，所以 Mean（Y1） = Mean（Time）；而 Mean（Y2） = L2 * Mean（Time） + Mean（e2），亦即 Mean（Y2） = L2 * Mean（Y1），移项

即可推演出 L2 之路径系数等于$\frac{\text{Mean(Y2)}}{\text{Mean(Y1)}}$,同理可推出 L3 之路径系数等于$\frac{\text{Mean(Y3)}}{\text{Mean(Y1)}}$, L4 之路径系数等于$\frac{\text{Mean(Y4)}}{\text{Mean(Y1)}}$,反映出 L2 与 L3 之回归系数是两个平均数值的比率。换言之,L1 是潜在特质的起点状态,以第一次测量结果为起点行为,经由此特殊设计,研究者可以探究后续之重复测量与第一次测量间之改变量比率,以判定成长之速率是逐渐增加或减少。

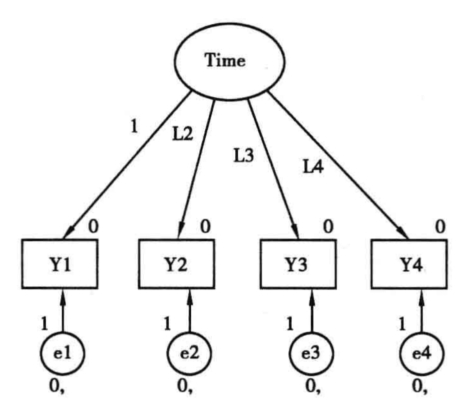


图 5-6　单因子 LCA 线性模型

图 5-7 中的协方差矩阵,是单因子 LCA 模型的数据文件,纪录了 118 位青少年在 12, 13,14,15 岁时之饮酒量,这个数据文件是利用 Excel 创建的,以供 Amos 参数估计时的调用。

嘉義大學國教所, Fred Li, 2007 - Lca1.xls

	A	B	C	D	E	F
1	rowtype_	varname_	Y1	Y2	Y3	Y4
2	n		118	118	118	118
3	cov	Y1	0.855			
4	cov	Y2	0.473	0.82		
5	cov	Y3	0.408	0.524	0.871	
6	cov	Y4	0.3	0.446	0.496	0.761
7	Mean		1.788	2.102	2.347	2.737

图 5-7　单因子 LCA 模型的数据文件

资料来源:Taken from http://osu. orst. edu/dept/hdfs/papers/paper. htm

根据相关之理论,研究者可以提出 LCA 误差独立模型与 LCA 误差相关模型,分别如图 5-8-1 与图 5-8-2 所述。统计分析结果亦如图所示。

由图 5-8-1 中单因子 LCA 独立模型的 Amos 分析结果得知,Level 的平均数为 1.80,表示这群少年在 12 岁时的平均饮酒期望值,其饮酒之变异量为 0.27;后续之重复测量与第一次测量间的改变量比率分别为 1.18, 1.31, 1.51,显示这批青少年的饮酒量逐年增加呈线性关系。这批 12 ~ 15 岁青少年饮酒的期望值为:

E(Y4)	E(Y3)	E(Y2)	E(Y1)
2.70570707	2.36146613	2.12418564	1.79598256

由前述公式知,上述期望值的计算公式为:

$$E(Y1) = 1 * E(Mean_{int}) \approx 1.8;$$

$$E(Y2) = 1.18 * E(Mean_{int}) \approx 1.18 * 1.80 \approx 2.1;$$

$$E(Y3) = 1.31 * E(Mean_{int}) \approx 1.31 * 1.80 \approx 2.4;$$

$$E(Y4) = 1.51 * E(Mean_{int}) \approx 1.51 * 1.80 \approx 2.7。$$

这些期望值可利用 Amos 属性窗口 Output 下的[Implied moment]按钮,轻松求得。

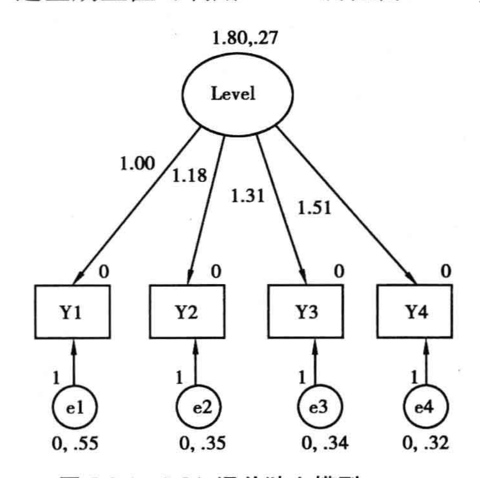


图 5-8-1 LCA 误差独立模型

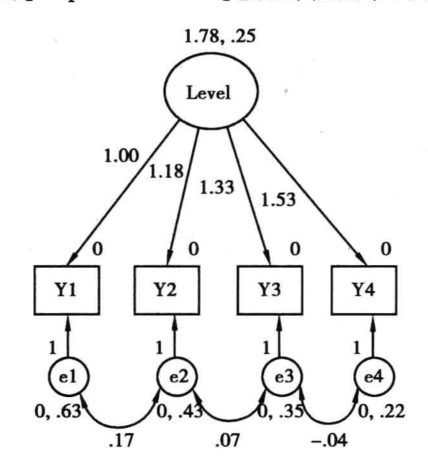


图 5-8-2 LCA 误差相关模型

另外,经适配度检验结果知,χ^2 = 17.379,df = 5,p =0.038,NFI =0.899,TLI = 0.91,RMSEA = 0.145,表示本单因子线性模型与数据适配的适配度,似乎未臻理想;但如将测量误差间的两两相关亦列入参数的估计(如图 5-8-2),其 χ^2 =5.7722, df = 2, p =0.0558,适配度似乎有明显的改善,这意味着重复测量间必然具有或多或少的相关性。事实上,在重复量数的测量上,目前的状态由过去所决定,而目前的状态又支配着未来的发展,必然存在着自相关。

另外,研究者如欲针对 Y1 ~ Y4 这几对理论平均数间的差异进行统计检验,可以利用 Amos 的模型管理窗口,进行路径系数间的等同限制设定。操作步骤简述如下:

1. 绘制如图 5-6 之路径图;
2. 打开图 5-9 功能表的 Analyze 表单,点选[Manage Models];
3. 在图 5-10 的 Mange Models 窗口中,进行模型的命名(L2 = L3 = L4)与相关参数的等同限制;
4. 重复步骤 3 的步骤,再建立 L2 = L3,L2 = L4,L3 = L4 等模型命名与相关参数的等同限制,以上五个理论模型的设定结果,Amos 会显示于如图 5-11 右下角的小窗口内。

前述五个理论模型的 Amos 分析结果,逐一列述如表 5-1 ~5-3。

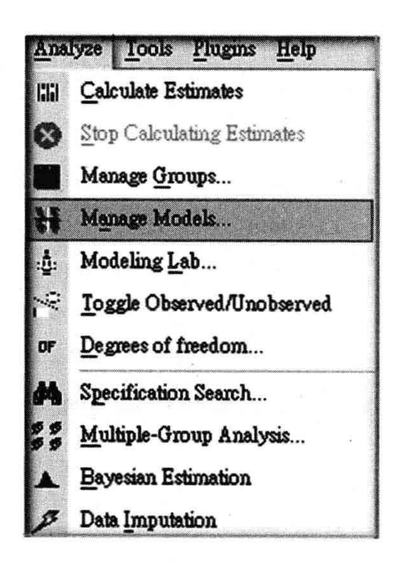


图 5-9　模型管理窗口点选表单

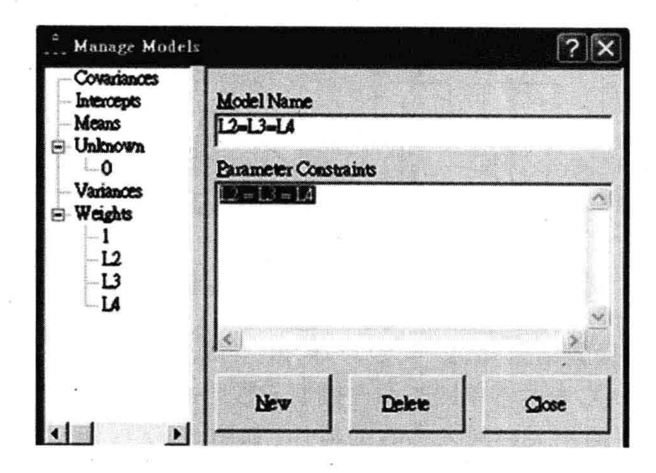


图 5-10　模型管理窗口

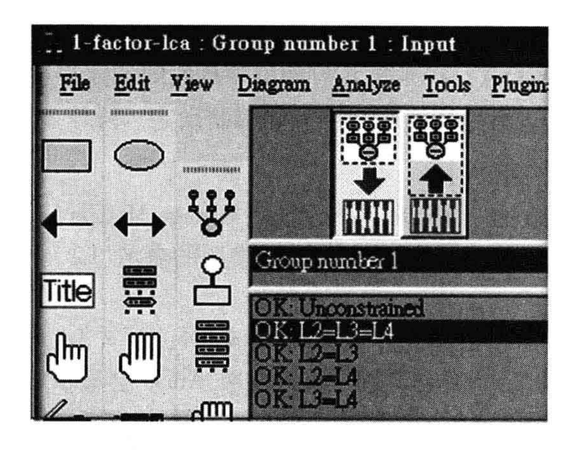


图 5-11　模型管理窗口的设定结果显示窗口

表 5-1 Amos 报表:Assuming model Unconstrained to be correct

Model	DF	CMIN	P	NFI Delta-1	IFI Delta-2	RFI rho-1	TLI rho2
L2 = L3 = L4	2	51.234	.0000	.299	.308	.221	.229
L2 = L3	1	9.851	.0016	.057	.059	.037	.038
L2 = L4	1	51.228	.0000	.298	.307	.278	.288
L3 = L4	1	20.283	.0000	.118	.121	.098	.101

表 5-2 Amos 报表:Model Fit Summary

Model	NPAR	CMIN	DF	P	CMIN/DF
Unconstrained	9	17.3788	5	.0038	3.475
L2 = L3 = L4	7	68.6134	7	.0000	9.801
L2 = L3	8	27.2304	6	.0001	4.538
L2 = L4	8	68.6074	6	.0000	11.434
L3 = L4	8	37.6626	6	.0000	6.277

由表5-1卡方检验($p<0.01$)与表5-3检验结果RMSEA(>0.05)来看,这五个理论模型的适配度似乎不佳,显示出在四次的测量上,其因素负荷量配对间具有显著差异。

表 5-3 Amos 报表:主要模型适配度指标

Model	NFI Delta1	RFI rho1	RMSEA	AIC	TLI rho2	CFI
Unconstrained	.8985	.8782	.1454	35.378	.910	.925
L2 = L3 = L4	.5995	.6567	.2742	82.613	.680	.627
L2 = L3	.8410	.8410	.1739	43.230	.871	.871
L2 = L4	.5995	.5995	.2986	84.607	.621	.621
L3 = L4	.7801	.7801	.2123	53.662	.808	.808

兹将以上这四个重复测量间平均数差异的检验结果,整理合并摘要如表5-4。表5-4中,χ^2,p与df值取自表5-2,而$\Delta\chi^2$,p与Δdf值取自表5-1。

表 5-4 各对重复测量间平均数差异检验摘要表

路径系数之等同限制	χ^2	df	$\Delta\chi^2$	ΔDF	p
未受限	17.379	5			
L2 = L3 = L4	68.613	7	51.235	2	.0000
L2 = L3	27.230	6	9.851	1	.0016
L2 = L4	68.607	6	51.229	1	.0000
L3 = L4	37.663	6	20.284	1	.0000

注:$\Delta\chi^2$系指各限制模型与未设限模型之比较。

由表 5-4 知，这四个路径系数相等的假设（相当于平均数相等的假设），均未获得支持（$p<0.01$）。由 L2 = L3 = L4 设限模型与未设限模型之 $\Delta\chi^2$（= 51.235）与 df = 2 来看，以第一次测量为基准点，后续三次测量的平均数改变的比率并不相同，其余各对平均数改变的比率亦均不相等（$L2 \neq L3, L2 \neq L4, L3 \neq L4$）。再根据前述这批 12～15 岁青少年饮酒的期望值，可获得以下结论：①13～15 岁青少年的饮酒量均显著与 12 岁时不同；②青少年饮酒量有逐年增加的趋势。

第五节　双因子潜在特质变化分析

双因子潜在特质变化分析涉及测量基线因子（level factor）的分析与改变型态因子（shape factor）的分析（邱皓政，2004）。这两个因素又称为起点状态因素及改变因素（余民宁，2006），因而除了可以分析个别重复测量变量在潜在特质上的初始状态之外，改变型态因子的加入还可让研究者去分析成长曲线是线性还是非线性、成长趋势是正向还是负向。改变型态的分析主要可分为线性、片段线性与非线性三种（参见图 5-12），检验的途径主要有二：①图示法；②一系列的适配度检验（Curran & Hussong, 2003）。

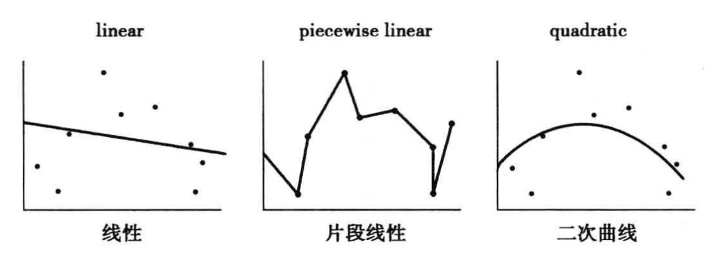


图 5-12　常见的改变型态模型

一、线性分析

因为初始状态因子（Level）类似于回归分析中的截距，这个因素上的四次测量的因素负荷量均固定为 1，代表起始水平。研究者如欲探究最后一次与第一次测量间的改变差异量及比较各个测量点间的发展比率，Shape 因子的时间编码须修改为如图 5-13 内之路径系数：0，L2，L3，1，其中 L2 与 L3 系表示开放估计，以了解成长之速率是逐渐增加或减少。图 5-13 二因子 LCA 模型中，SHAPE 的 Y1 路径系数设定为 0，乃是希望将时间的原点放在 12 岁作为基准点。因此，图 5-14 中 Level 的平均值代表青少年在 12 岁时的平均饮酒量（经测量误差调整）。此外，Level 与 Shape 间具有共变关系，表示起始水平与其后线性改变间的相关。至于 Level 与 Shape 的变异量，则是代表起始水平与其后线性改变量的个别差异情形，亦反映出这个模型是否适用于大多数的个案。当然研究者如欲分析误差间是否具有相关，亦可在图 5-13 路径图中的 E1～E4 间加上共变关系，这是传统 ANOVA 分析无法做到的。

由图 5-13 知，此二因子 LCA 模型的测量回归模型可陈述如下：

$$Y1 = 1.0 * Level + 0 * Shape + e1;$$
$$Y2 = 1.0 * Level + L2 * Shape + e2;$$
$$Y3 = 1.0 * Level + L3 * Shape + e3;$$
$$Y4 = 1.0 * Level + 1.0 * Shape + e4。$$

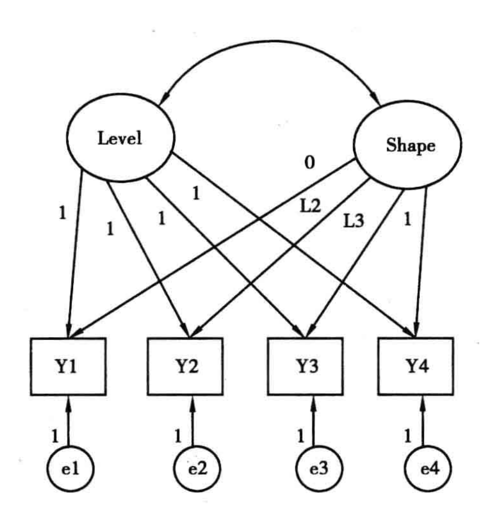


图 5-13　二因子 LCA 模型

可将上列式子化成矩阵方式呈现：$\begin{bmatrix} Y1 \\ Y2 \\ Y3 \\ Y4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & L2 \\ 1 & L3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} Level \\ Shape \end{bmatrix} + \begin{bmatrix} e1 \\ e2 \\ e3 \\ e4 \end{bmatrix}$，式中 Level 系初始状态因子，Shape 系变化比率因子。利用前述之方差、协方差与平均数的运算原则 5，可推得：Mean(Y1) = 1.0 * Mean(Level) + Mean(e1)，因为 Mean(e1) = 0，所以 Mean(Y1) = Mean(Level)；而 Mean(Y4) = 1.0 * Mean(Level) + 1.0 * Mean(Shape) + Mean(e 4)，亦即 Mean(Y4) = Mean(Y1) + Mean(Shape)，移项 Mean(Shape) = Mean(Y4) - Mean(Y1)。有了 Mean(Y1) = Mean(Level) 与 Mean(Shape) = Mean(Y4) - Mean(Y1) 的关系公式，再利用 Mean(Y2) = 1.0 * Mean(Level) + L2 * Mean(Shape) + Mean(e2)，即可推演出 L2 的路径系数等于$\frac{Mean(Y2) - Mean(Y1)}{(Mean(Y4) - Mean(Y1))}$；同理，L3 之路径系数等于$\frac{Mean(Y3) - Mean(Y1)}{(Mean(Y4) - Mean(Y1))}$，反映出 L2 & L3 的回归系数系两个平均数差异值(gain score or change score)的比率，如此设计可以有效地将初始效应从后续之特质改变量中分离开来。换言之，Level 系潜在特质的起点状态，以第一次测量结果为起点行为，而 Shape 是潜在特质在终点时的改变状态，以第四次测量结果为终点行为。经由此特殊设计，研究者可以探究后续的重复测量与第一次测量间的改变量比率，并判定成长的速率是逐渐增加还是减少。

二因子 LCA 模型的分析，仍延用图 5-7 的资料，该文件是 118 位青少年在 12，13，14，15 岁时的饮酒量，这个数据文件利用 Excel 建立，以便 Amos 调用。将图 5-13 的二因子 LCA 模型与图 5-7 数据文件相互联结后，利用 Amos 分析的结果，如图 5-14 所示。

图 5-14 中系数是二因子 LCA 模型 Amos 分析结果，Level 的平均数为 1.79，代表这群少年在 12 岁时的平均饮酒量，其方差为 0.55；而 SHAPE 的平均数为 0.95，表 12 岁与 15 岁青少年饮酒的平均差异量，因为 Mean(Shape) = Mean(Y4) - Mean(Y1) = 2.74 - 1.79 = 0.95；Shape 的路径系数分别为 0，0.34，0.59，1，亦反映出饮酒量与年龄之间的线性关系，显示这批青少年饮酒量逐年增加。这批 12 ~ 15 岁青少年饮酒的期望值为：

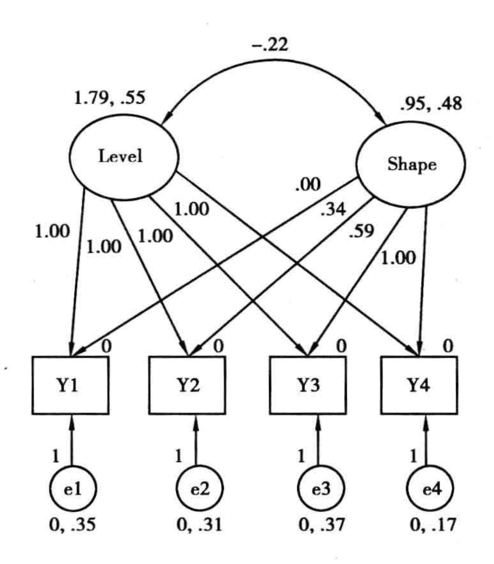


图 5-14 二因子 LCA 模型 Amos 分析结果

Y4	Y3	Y2	Y1
2.73705688	2.34250719	2.10812658	1.78516002

上述期望值的计算公式为:E(Y1) = 1 * E($Mean_{int}$) + E($Mean_{shape}$)At≈1.79 +0.95 * 0≈1.79,以此类推,E(Y2) = 1.79 + 0.95 * 0.34≈2.108,E(Y3) = 1.79 + 0.95 * 0.59 ≈2.34,E(Y4) = 1.79 +0.95 * 1≈2.737。

Y1 ~ Y4 的期望平均数如图 5-15,亦反映出饮酒量与年龄之线性关系。

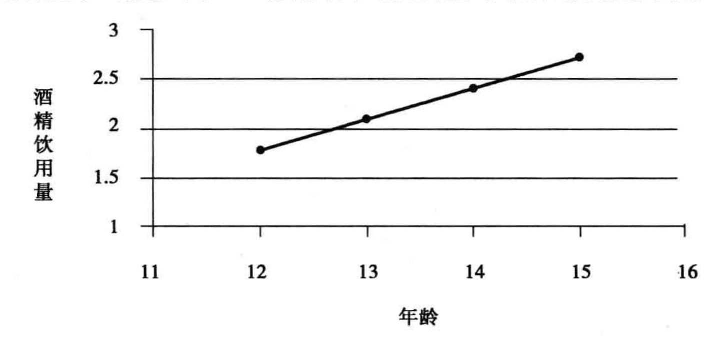


图 5-15 饮酒量与年龄的线性关系

另外,根据 χ^2 =2.46,df =3,p =0.4832,NFI =0.9857,CFI =1.0000,RMSEA =0.0000,亦表二因子线性模型与数据相当适配。

研究者如欲探究在各年龄间(12 ~ 15)饮酒量的平均改变量,Shape 相关之时间编码须修改为如图 5-16 内之线性路径系数:0, 1, 2, 3。

利用前述之方差、协方差与平均数的运算原则 5,可推得:Mean(Y1) = 1.0 * Mean(Level) + 0 * Mean(Shape) + Mean(e1),因为 Mean(e1) = 0,所以 Mean(Y1) = Mean(Level),亦即 Level 的平均数等于 Y1 的期望值;而 Mean(Y2) = 1.0 * Mean(Level) + 1.0 * Mean(Shape) + Mean(e2),亦即 Mean(Y2) = 1 * Mean(Y1) + 1 * Mean(Shape),移

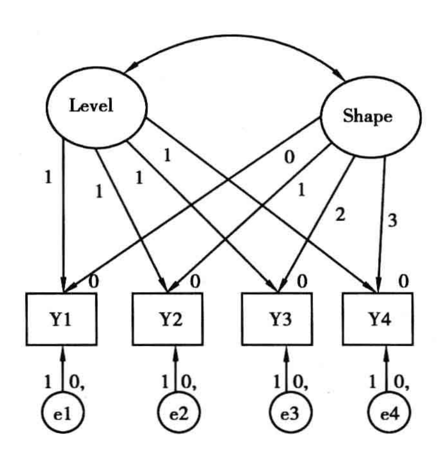


图 5-16　二因子 LCA 线性模型

项 Mean(Shape) = (Mean(Y2) - Mean(Y1))/1；Mean(Y3) = 1.0 * Mean(Level) + 2.0 * Mean(Shape) + Mean(e3)，亦即 Mean(Y3) = Mean(Y1) + 2.0 * Mean(Shape)，Mean(Y4) = 1.0 * Mean(Level) + 3.0 * Mean(Shape) + Mean(e4)，亦即 Mean(Y4) = Mean(Y1) + 3.0 * Mean(Shape)。同理移项类推，Mean(Shape) = (Mean(Y2) - Mean(Y1))/1 = (Mean(Y3) - Mean(Y1))/2 = (Mean(Y4) - Mean(Y1))/3(假如图 5-16 的理论模型正确无误的话)。由此可知，Shape 的平均值反映出在各年龄间饮酒量的平均改变量(=0.31，以第一次做比较的平均差异)，大约以 0.31 的速度增加，而此改变差异的变异量(方差)为 0.05，显示个别间的改变量差异不大，此成长模型适用于所有受试者，请参阅图 5-17-1 中 Shape 的平均数(31)与方差(0.05)。在图 5-17-2 中 Level 与 Shape 的相关系数为 -0.41，整体而言初期的饮酒量愈高者，其后续几年的饮酒改变比率似乎较小，呈现负相关现象。

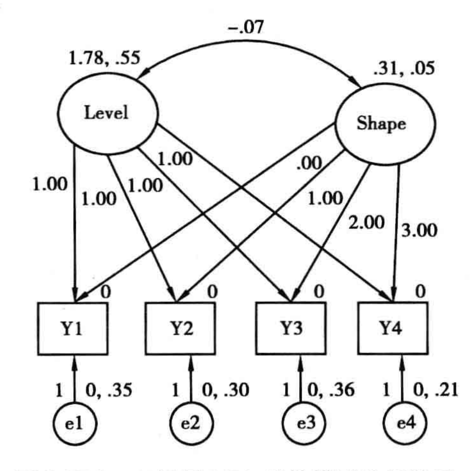


图 5-17-1　二因子 LCA 线性模型分析结果

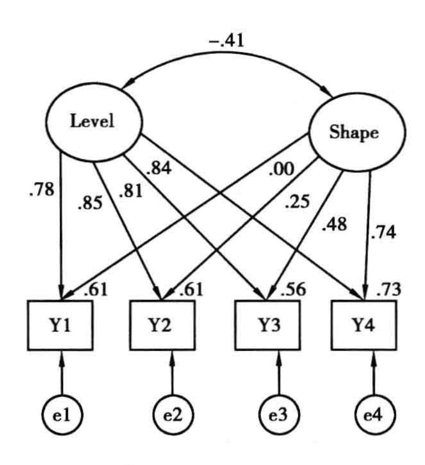


图 5-17-2　二因子 LCA 线性模型分析结果(标准化)

至于个体的线性潜在特质在某一时间点上的变化分析 Y_{it}，则须借助于下列 3 个个体回归方程式：

$$Y_{it} = a_t + \lambda_0 Level_i + \lambda_t Shape_i + \varepsilon_{it}$$

$$Level_i = \overline{Level} + \zeta_{li}$$

$$Shape_i = \overline{Shape} + \zeta_{si}$$

式中 λ_0 系起点状态因子 Level 的 basis function（改变轨迹的基本形式），全部设定为 1，λ_t 系型态因子 Shape 的 basis function，α_t 是各指标的截距，因每一时间点上仅有一个指标，全部设定为 0，ζ_{1i} 与 ζ_{si} 是个体与 Level 及 Shape 平均数的离均差。当整体性的 LCA 分析不适配时，个体的线性潜在特质在某一时间点上的变化分析，就显得格外的重要。

二、片段线性成长模型

片段线性成长模型（Piecewise growth modeling），外观看起来像曲线成长轨迹，但它可被分割成几个线性成长区段，我们称之为片段线性成长模型。假设有一研究者欲探讨小学 1～5 年级学生的学习压力成长趋势，经实际观察发现小学生的学习压力从 1～3 年级呈现缓慢的成长，但从 3～5 年级则呈现陡升的现象。因此，该研究者提出图 5-18 的片段线性成长模型之路径图（改编自 MPLUS 范例 6-11），以检验本模型的适配度。

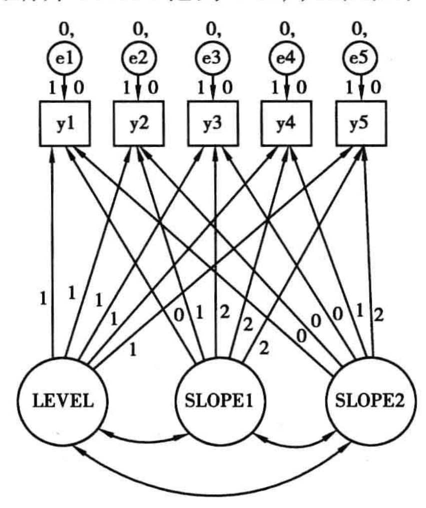


图 5-18 片段线性成长模型路径图

利用前述方差、协方差与平均数的运算原则 5，可推得：Mean（Y1） = 1.0 * Mean（Level） + 0 * Mean（Slope1） + 0 * Mean（Slope2） + Mean（e1），因为 Mean（e1） = 0，所以 Mean（Y1） = Mean（Level），亦即 Level 的平均数等于 Y1 的期望值；而 Mean（Y2） = 1.0 * Mean（Level） + 1 * Mean（Slope1） + 0 * Mean（Slope2） + Mean（e2），亦即 Mean（Y2） = 1 * Mean（Y1） + 1 * Mean（Slope1），移项 Mean（Slope1） = （Mean（Y2） − Mean（Y1））/1；Mean（Y3） = 1.0 * Mean（Level） + 2 * Mean（Slope1） + 0 * Mean（Slope2） + Mean（e3），亦即 Mean（Y3） = Mean（Y1） + 2.0 * Mean（Slope1）；Mean（Y4） = 1.0 * Mean（Level） + 2 * Mean（Slope1） + 1 * Mean（Slope2） + Mean（e4），亦即 Mean（Y4） = Mean（Y1） + 2 * Mean（Slope1） + 1 * Mean（Slope2）；Mean（Y5） = 1.0 * Mean（Level） + 2 * Mean（Slope1） + 2 * Mean（Slope2） + Mean（e4），亦即 Mean（Y5） = Mean（Y1） +2 * Mean（Slope1） + 2 * Mean（Slope2）。研究者利用以上公式，即可求得 Y1～Y4 的期望值。

由图 5-19 知，小学在一年级时，学习压力之平均数为 0.46，方差为 0.98，反映出小学生刚入学时之学习压力个别差异很大。再由 Slope1 的平均数为 1.07 爬升到 Slope2 的平

均数 1.96 来看,显现出后期小学生压力陡升的现象。

再由图 5-20 及表 5-5 中的适配度指标来看,本模型是非常适配的(p = 0.514, NFI = 0.997, RMSEA = 0.0000)。

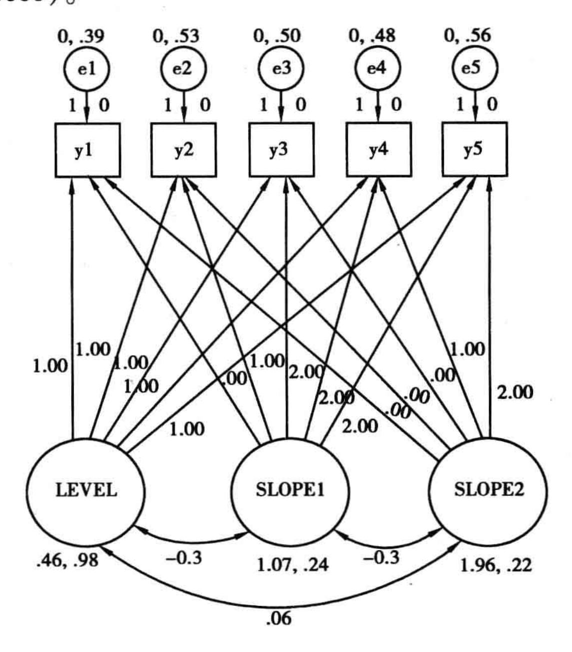


图 5-19　片段线性成长模型分析结果(未标准化)

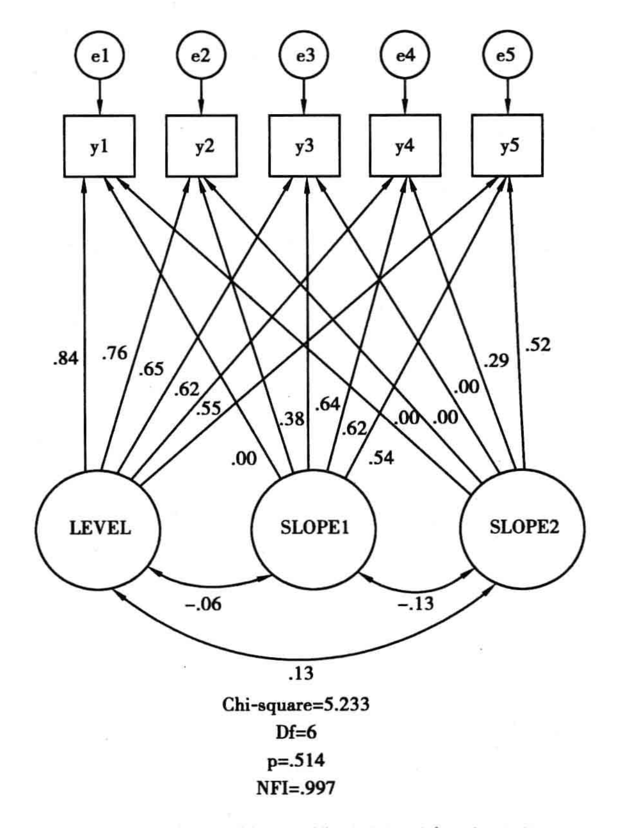


图 5-20　片段线性成长模型分析结果(标准化)

假如模型适配度佳，则上述$\overline{Y}_1 \sim \overline{Y}_5$ 的平均数期望值(计算方法请参照前节说明)：0.4622，1.5336，2.6050，4.5623，6.5196，会接近于实际观察值：0.4432，1.5841，2.5949，4.5354，6.5351。兹将 Amos 的统计分析报表，摘要如表 5-5 ~ 表 5-10。研究者可以利用平均数期望值，绘制如图 5-21 片段线性成长曲线，在小学 3 年级时，曲线出现转折的现象。图 5-21 成长曲线亦反映出：小学生之学习压力从 1 ~3 年级呈现缓慢的成长，从3 ~5 年级则呈现陡升的现象。

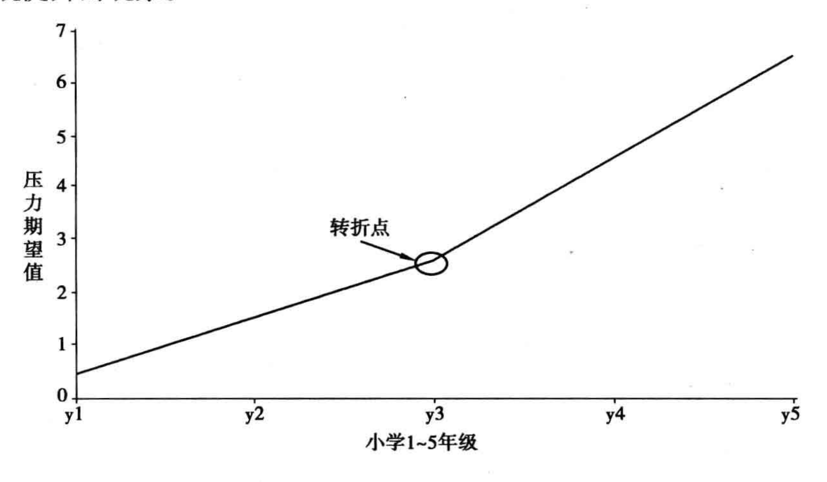


图 5-21 片段线性成长曲线

表 5-5 Baseline Comparisons

Model	NFI Delta1	RFI rho1	RMSEA	IFI Delta2	TLI rho2	CFI
Default model	.9967	.9945	.0000	1.0005	1.0008	1.0000

表 5-6 Standardized Regression Weights

			Estimate
y1	<---	LEVEL	.8450
y2	<---	LEVEL	.7630
y3	<---	LEVEL	.6504
y4	<---	LEVEL	.6246
y5	<---	LEVEL	.5503
y1	<---	SLOPE2	.0000
y5	<---	SLOPE1	.5429
y4	<---	SLOPE1	.6162
y3	<---	SLOPE1	.6416
y2	<---	SLOPE1	.3764
y1	<---	SLOPE1	.0000
y2	<---	SLOPE2	.0000
y3	<---	SLOPE2	.0000
y4	<---	SLOPE2	.2946
y5	<---	SLOPE2	.5191

表 5-7 Means

	Estimate	S. E.	C. R.	P
LEVEL	.4622	.0515	8.9674	* * *
SLOPE1	1.0714	.0298	35.9804	* * *
SLOPE2	1.9573	.0304	64.3674	* * *

表 5-8 Covariances

			Estimate	S. E.	C. R.	P
LEVEL	<-->	SLOPE1	−.0291	.0492	−.5917	.5540
SLOPE2	<-->	SLOPE1	−.0309	.0261	−1.1851	.2360
LEVEL	<-->	SLOPE2	.0594	.0362	1.6427	.1004

表 5-9 Correlations

			Estimate
LEVEL	<-->	SLOPE1	−.0599
SLOPE2	<-->	SLOPE1	−.1349
LEVEL	<-->	SLOPE2	.1279

表 5-10 Variances

	Estimate	S. E.	C. R.	P
LEVEL	.9846	.1011	9.7376	* * *
SLOPE2	.2191	.0417	5.2520	* * *
SLOPE1	.2396	.0368	6.5038	* * *
E1	.3944	.0760	5.1897	* * *
E2	.5251	.0431	12.1787	* * *
E3	.5013	.0675	7.4232	* * *
E4	.4832	.0478	10.1036	* * *
E5	.5587	.1035	5.3959	* * *

三、非线性分析

虽然当时间间距较短时,线性模型的分析较精简,但当时间间距较长时,必须使用非线性模型来正确描述成长之轨迹。非线性模型常见的有二次函数(quadratic function)与指数函数(exponential function)。以下将以高三学生联考前四个月的压力指数的例子,介绍如何利用 Amos 来进行二次多项式(Quadratic polynomial)的非线性潜在特质变化分析。本实例的数据是 500 位高三学生联考前四个月(Y1 ~ Y4)的压力指数的平均数、协方差

矩阵(参见图 5-22)。

	rowtype_	varname_	y1	y2	y3	y4
1	n		500.0000	500.0000	500.0000	500.0000
2	cov	Y1	1.9355	.	.	.
3	cov	Y2	1.1762	2.9113	.	.
4	cov	Y3	1.3057	3.2881	7.9597	.
5	cov	Y4	1.6380	4.9559	12.4795	24.9473
6	Mean		.5164	2.0875	4.6185	8.2396

图 5-22　高三学生联考前四个月的压力指数的平均数、协方差矩阵

图 5-23 之路径图设计,是二因子 LCA 的非线性模型,图中 Level 因子是截距,所有相关之路径系数均设定为 1,表示任何受试者在四次的重复测量上,截距均为常数固定值。图中 Slope1 因子是线性斜率,代表与每一重复测量相关的时间,线性斜率被设定为 0,表示测量之起点。图中 Slope2 因子是二次曲线斜率,为线性斜率的平方值(参见图 5-26 中参数限制的设定)。

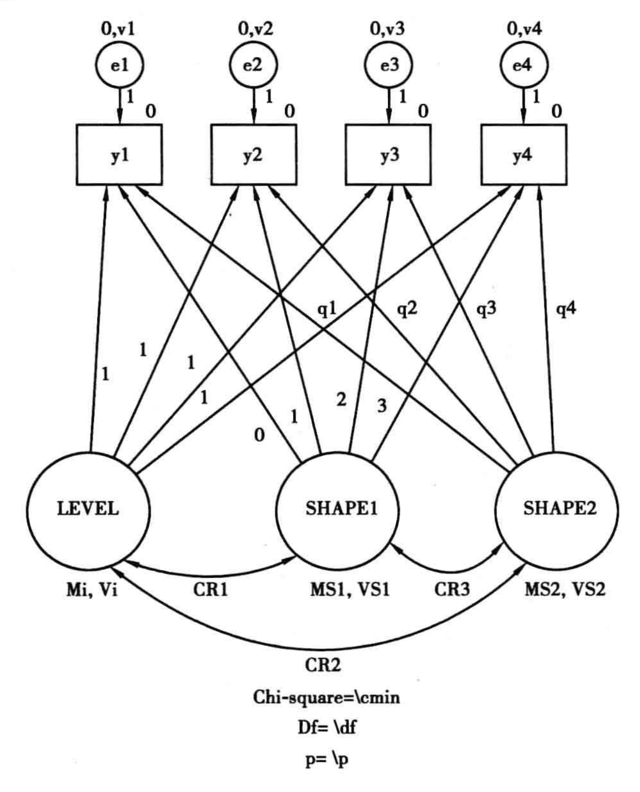


图 5-23　二因子 LCA 非线性模型

为进行线性模型与非线性模型间的统计检验,研究者可借用图 5-25 的 Amos 的模型管理窗口,让 Amos 自动为您进行 χ^2 差异值的检验。

图 5-24 右侧中的 Linear model 与 Quadratic model 的设定,是利用 Amos 中图 5-25 与图 5-26 的管理窗口,在该窗口中参数受限的方框中,进行相关参数的设定。在 Linear model 中,必须将 Quadratic model 中的相关参数全部设定为 0(如斜率 q1 ~ q4)。

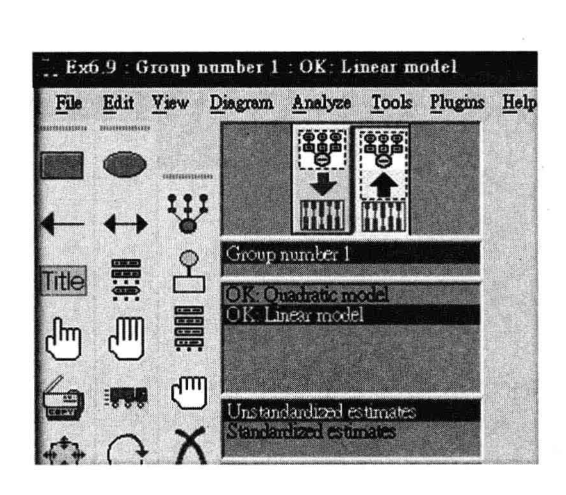


图 5-24　Amos 之 LCA 线性与非线性模型的设定

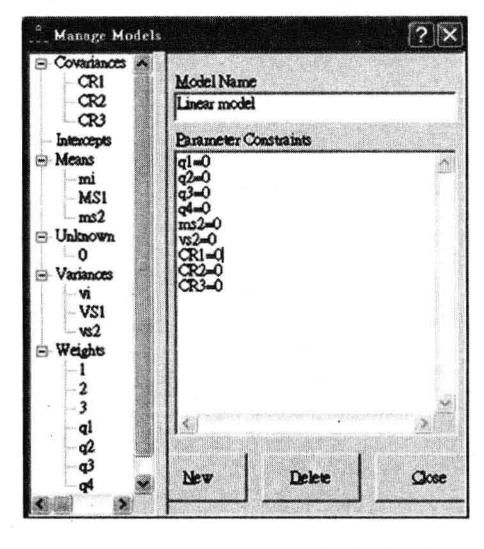


图 5-25　Amos 之 LCA 线性模型设定

在 Quadratic model 中,必须将 Quadratic model 中的斜率参数(q1 ~ q4)设定为 0, 1, 4, 9,是 Linear model 中斜率 0, 1, 2, 3 的平方值,参见图 5-26。

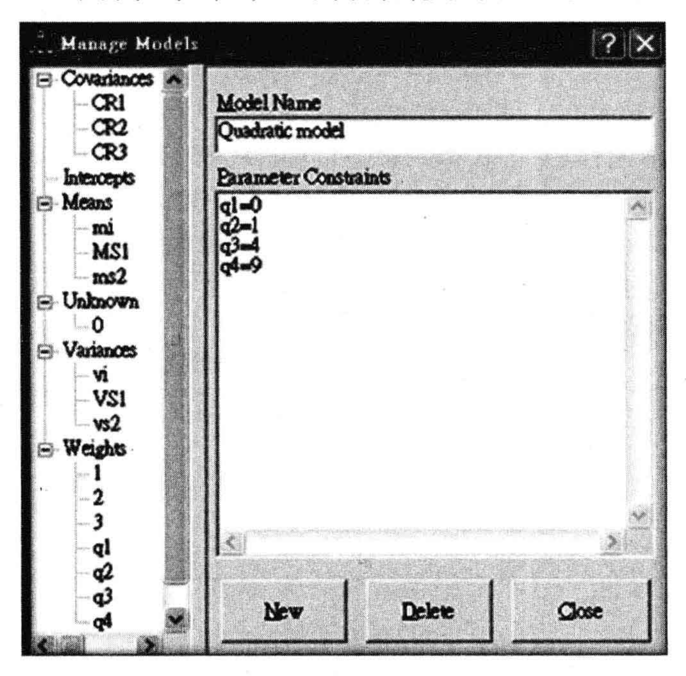


图 5-26　Amos 之 LCA 线性二次方模型设定

由图 5-27 的分析结果知,本实例线性分析时,出现 Y3 的误差方差为负(-0. 17)的情形,此乃模型界定错误所造成的非正定矩阵的实例。因此,非正定矩阵的出现,可能意

味着这些高中学生的前压力成长变化具有非线性的趋势。根据 χ^2 = 333.158, df = 6, p = 0.000,亦反映线性模型与数据的不适配(注意,非正定矩阵的参数与适配指标,因无法正确估计在此不列出)。

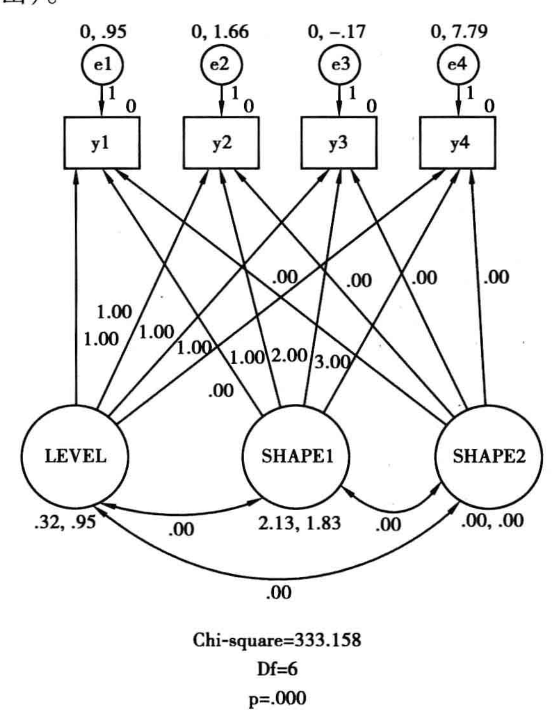


图 5-27　二因子 LCA 线性模型 Amos 分析结果

由图 5-28 之二因子 LCA 非线性模型 Amos 分析结果知,本数据的二次多项式模型与数据具有极佳的适配性(χ^2 = 0.472, df = 1, p = 0.492, NFI = 0.9996, GFI = 1.000, RMSEA = 0.0000),非正定矩阵的现象亦不复见,反映出模型界定正确的重要性。

根据 Amos 的报表,以下将线性与非线性模型差异检验结果摘要如表 5-11。

由表 5-11 知,非线性与线性模型间具有显著差异($p < 0.01$),显示出对于高中生,联考压力与距离联考的时间具有二次曲线关系。

表 5-11　线性与非线性模型差异检验摘要表

模　型	χ^2	df	ΔDF	p
线性	333.158	6		.0000
非线性	.472	1		.492
$\Delta\chi^2$	332.686		5	.0000

表 5-12　Amos 报表:Model Fit Summary

Model	NFI Delta1	RFI rho1	RMSEA	TLI rho2	CFI
Quadratic model	.9996	.9977	.0000	1.0026	1.0000
Linear model	.7289	.7289	.3306	.7325	.7325

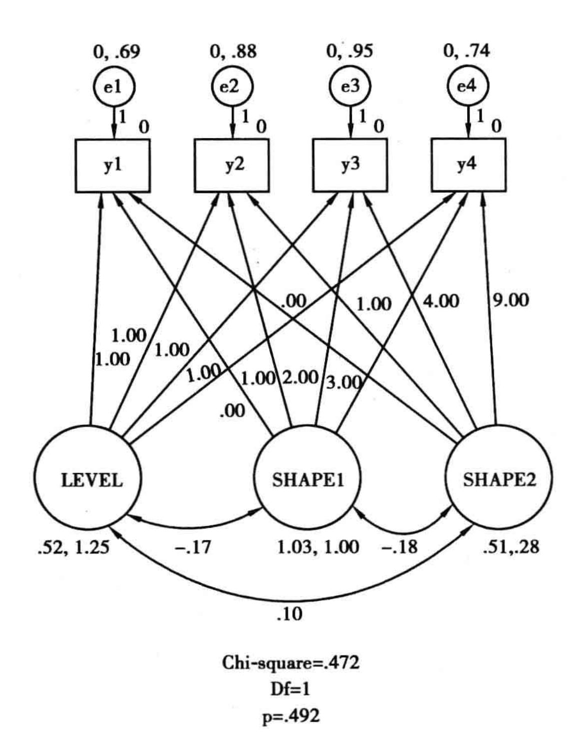


图 5-28　二因子 LCA 非线性模型 Amos 分析结果

表 5-13　Linear model 之期望值

E(Y4)	E(Y3)	E(Y2)	E(Y1)
6.7212	4.5873	2.4534	.3195

另外,根据图 5-27 中相关统计量,上述线性期望值的计算公式为: $E(Y1)=1*E(Mean_{int})+E(Mean_{linear})A_t\approx0.32$,以此类推, $E(Y2)=0.32+2.13*1\approx2.45$, $E(Y3)=0.32+2.13*2\approx4.58$, $E(Y4)=0.32+2.13*3\approx6.72$。

表 5-14　Quadratic model 之期望值

E(Y4)	E(Y3)	E(Y2)	E(Y1)
8.2343	4.6392	2.0682	.5214

再根据图 5-28 中得相关统计量,上述非线性期望值的计算公式为: $E(Y1)=1*E(Mean_{int})+E(Mean_{shape})A_t+E(Mean_{Quadratic})(A_t)^2\approx0.52+1.03*0+0.51*0^2\approx0.52$,以此类推, $E(Y2)=0.52+1.03*1+0.51*1^2\approx2.06$, $E(Y3)=0.52+1.03*2+0.51*2^2\approx4.63$, $E(Y4)=0.52+1.03*3+0.51*3^3\approx8.2$。

根据表 5-13 与表 5-14 的期望平均数,绘制如图 5-29 的线形图,此图亦反映出二次曲线模型与线性模型间具有明显的差异,此趋势图正反映出前面统计检验的结果:二次曲线模型较符合高中生联考压力的成长趋势。

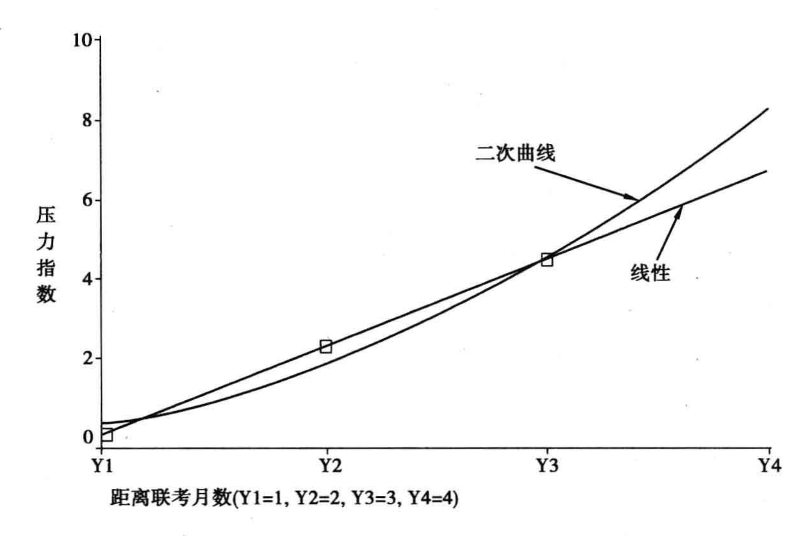


图 5-29　高中学生联考前压力与距离联考时间的相关曲线图

第六节　双因子潜在特质变化的共变分析

前述潜在特质变化模型成立的话,研究者亦可加入调节变量或协方差,以更进一步探讨有什么变量或因素会影响潜在特质的成长,这是一种有条件的潜在特质变化之共变分析(Conditional latent change modeling)。此种设计可以让研究者决定哪些变量会影响个别成长曲线的截距与斜率。此类潜在特质变化分析称为潜在特质变化的共变分析。图 5-30 系 275 位青少年连续四年(Y1 ~ Y4)接受每月吸烟量的测量的协方差矩阵,第一次测量时该批青少年从 13 ~ 16 岁都有,当第四次吸烟测量时,该批青少年的年纪介于 17 ~20 岁。吸烟量的测量采用五点量尺,1 表从未吸烟,5 表每天抽 1 根以上。而性别编码,男生登录为 1,女生登录为 0。为符合线性关系,图 5-30 之协方差矩阵中 age 与 gender 的协方差与方差经对数转换而得。

	A	B	C	D	E	F	G	H
1	rowtype_	varname_	Y1	Y2	Y3	Y4	gender	age
2	n		275	275	275	275	275	275
3	cov	Y1	1.776					
4	cov	Y2	1.624	1.98				
5	cov	Y3	1.451	1.604	1.885			
6	cov	Y4	1.265	1.46	1.556	1.899		
7	cov	gender	-0.0373	-0.00897	-0.00982	-0.03554	0.245	
8	cov	age	0.47	0.452	0.404	0.308	-0.0347	1.068
9	mean		2.2582	2.4655	2.5273	2.6945	0.4255	0.14236

图 5-30　潜在特质变化的共变分析数据

资料来源:Taken from http://osu. orst. edu/dept/hdfs/papers/paper. htm

在本例中,研究者不仅对青少年在 13 ~ 16 岁吸烟行为的趋势感兴趣,对于影响吸烟

行为的因素亦深感兴趣。因此,研究者提出性别与年龄两个相关因子,将这个因素加入潜在特质成长模型中,参见图 5-31 的路径设计。加入这两个预测变量的原因有二:

第一,研究者怀疑女性在初期较少吸烟,而导致后期比男性低的吸烟成长率。

第二,研究者怀疑在研究的初期,年纪较大者的吸烟量本来就比较大,但对后期吸烟行为并无影响。

利用前述之方差、协方差与平均数的运算原则 5,可推得:Mean(Y1) = 1.0 * [Mean(Level) + G0 * Mean(Gender) + A0 * Mean(Age)] + 0 * [Mean(Shape) + G1 * Mean(Gender) + A1 * Mean(Age)] + Mean(e1),因为 Mean(e1) = 0,所以 Mean(Y1) = 1.0 * [Mean(Level) + G0 * Mean(Gender) + A0 * Mean(Age)]。以此类推:Mean(Y2) = 1.0 * [Mean(Level) + G0 * Mean(Gender) + A0 * Mean(Age)] + 1.0 * [Mean(Shape) + G1 * Mean(Gender) + A1 * Mean(Age)] + Mean(e1);Mean(Y3) = 1.0 * [Mean(Level) + G0 * Mean(Gender) + A0 * Mean(Age)] + 2.0 * [Mean(Shape) + G1 * Mean(Gender) + A1 * Mean(Age)] + Mean(e1);Mean(Y4) = 1.0 * [Mean(Level) + G0 * Mean(Gender) + A0 * Mean(Age)] + 3.0 * [Mean(Shape) + G1 * Mean(Gender) + A1 * Mean(Age)] + Mean(e1)。研究者利用以上公式,即可求得 Y1 ~ Y4 的期望值。

图 5-31 的模型中,E5 是性别的变异量,E6 是年龄的变异量,并假设这两者之间具有相关。另外,性别与年龄变量加入模型中作为截距与斜率的预测指标,而且这两个预测变量是不随时间改变或改变缓慢的协方差(time-invariantcovariates)。此模型为二层次回归分析,第一阶是 Level 与 Shape 对指标变量(Y1 ~ Y4)的回归分析,第二阶是 Gender 与 Age 对潜在变量(Level 与 Shape)的回归分析。本模型假设这两个协方差 Gender 与 Age,通过潜在变量间接影响观察指标,亦即假定协方差与观察指标间的直接效果为 0,完全为中介效果。因此,图 5-31 的模型乃是一种成长预测模型。

由图 5-32 知,该批青少年初期的平均吸烟量为 2.23 根/每月,男生不管在初期的吸烟量还是后期的吸烟量改变比率均与女生无异(λ = -0.0643, 0.0077 均未达0.05 显著差异水平,p 值分别为 0.6743, 0.8552,参见表 5-19),每年平均以 0.15 根增加。

Y4	Y3	Y2	Y1
2.6949	2.5526	2.4102	2.2678

因而性别并不是影响青少年吸烟行为的主要因素。不过,年龄对于截距(Level)的影响力甚大(λ =0.4457, p =0.000),即每增加一岁就多抽 0.45 根香烟,反映出第一次测量时,年龄愈大的青少年吸烟量愈大,与其后的吸烟改变量亦具有显著关系,但效果值不大(λ = -0.0459, p =0.0234)。从图 5-33 中的标准化系数相对大小可知,性别不是吸烟量的重要影响因素,而年龄对于初测吸烟量较具预测力。表 5-15 到表 5-20 是 Amos 的统计分析结果,摘述于后供参考。

表 5-15 潜在变量期望值

	Estimate	S. E.	C. R.	P
Level	2.2317	.1013	22.0384	* * *
Shape	.1456	.0278	5.2344	* * *

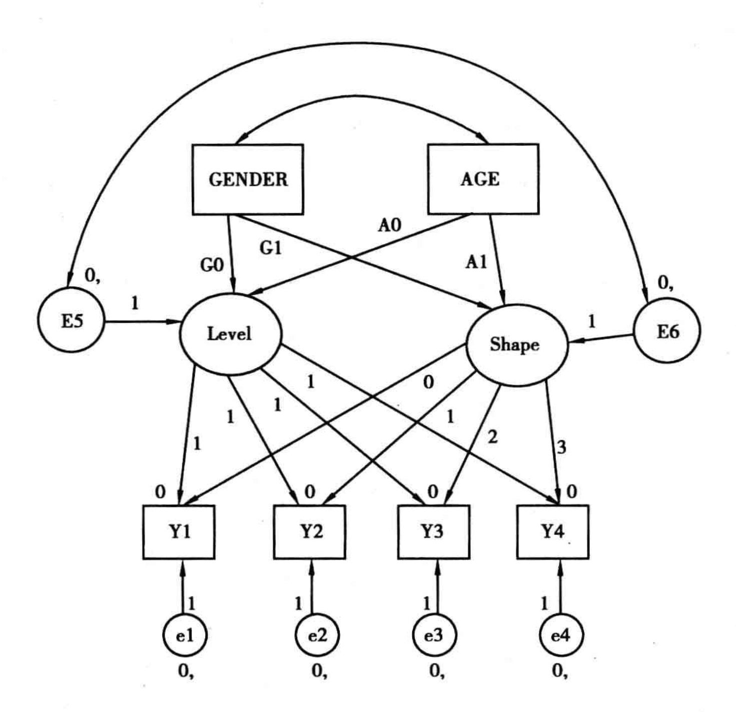


图 5-31 潜在特质变化的协方差分析路径图

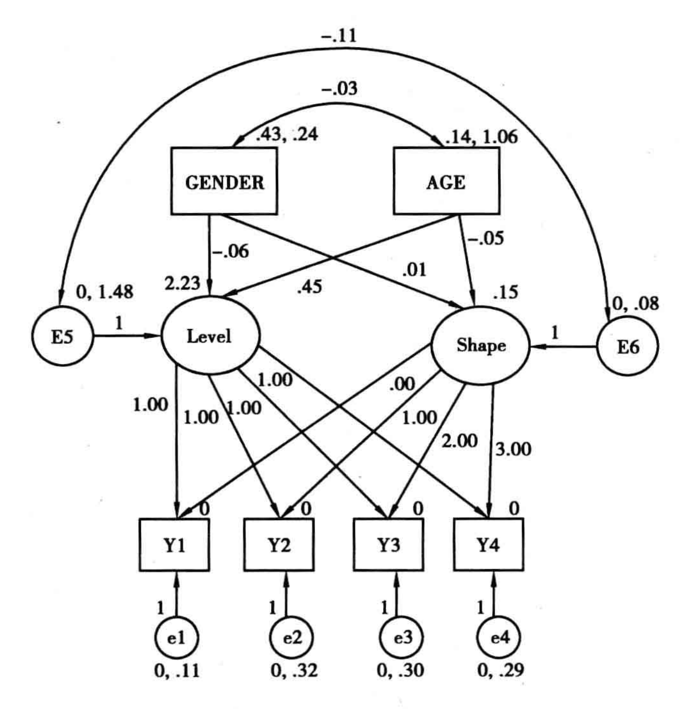


图 5-32 潜在特质变化的协方差分析结果

表 5-16　协方差估计值

			Estimate	S. E.	C. R.	P
AGE	<-->	GENDER	-.0346	.0309	-1.1203	.2626
E5	<-->	E6	-.1053	.0314	-3.3531	* * *

表 5-17　相关系数估计值

			Estimate
AGE	<-->	GENDER	-.0678
E5	<-->	E6	-.3037

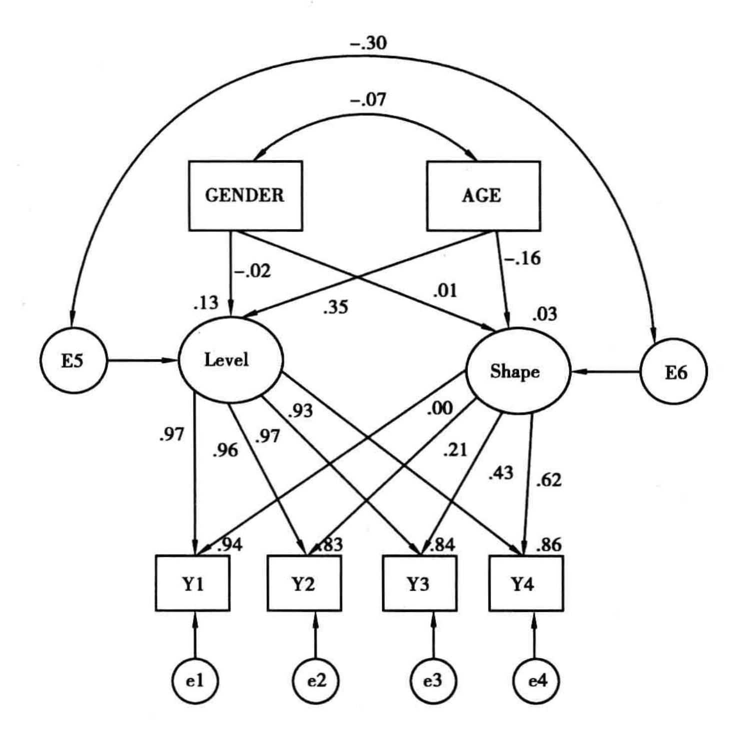


图 5-33　潜在特质变化之协方差分析结果(标准化)

表 5-18　方差估计值

	Estimate	S. E.	C. R.	P
AGE	1.0641	.0909	11.7047	* * *
GENDER	.2441	.0209	11.7047	* * *
E5	1.4826	.1391	10.6606	* * *
E6	.0812	.0128	6.3239	* * *
e1	.1100	.0456	2.4099	.0160
e2	.3190	.0358	8.9224	* * *
e3	.2990	.0358	8.3638	* * *
e4	.2854	.0573	4.9769	* * *

表 5-19　回归系数估计值

	Estimate	S. E.	C. R.	P	Label
Level <--- GENDER	−.0643	.1540	−.4176	.6763	G0
Shape <--- GENDER	.0077	.0423	.1825	.8552	G1
Level <--- AGE	.4457	.0737	6.0447	* * *	A0
Shape <--- AGE	−.0459	.0203	−2.2666	.0234	A1

表 5-20　模型适配度摘要表

Model	NPAR	CMIN	DF	P	NFI Delta1	RFI rho1	RMSEA	TLI rho2	PCFI	CFI
Default model	18	13.8071	9	.1294	.9873	.9788	.0442	.9925	.5973	.9955

假如协方差与观察指标间具有直接效果，而且研究的旨趣在于获得一个无偏估计值的模型，Stoel，Wittenboer 与 Hox（2004）推荐使用直接效果模型，尤其当前述的成长预测模型被推翻时，以避免偏差的统计估计值、模型界定错误或无法收敛的问题。此直接效果模型如图 5-34。

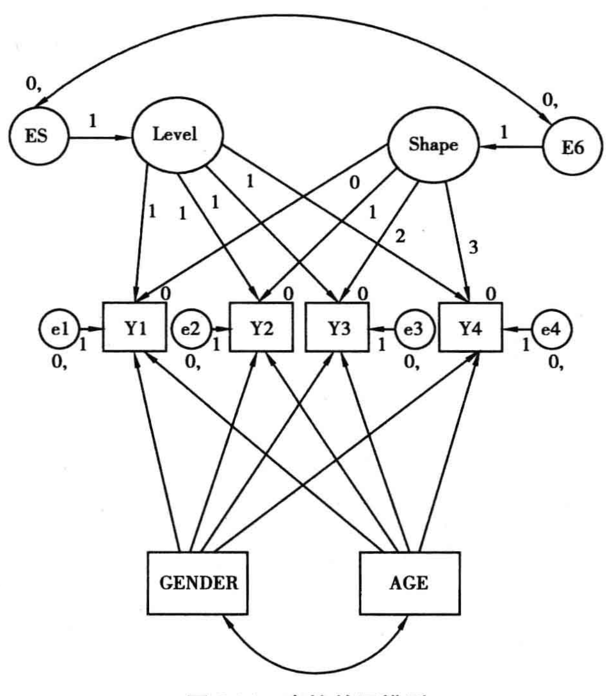


图 5-34　直接效果模型

综上所述，在进行成长曲线分析时，时间编码关系着相关路径系数或时间潜在因素平均数的解释，研究者须配合自己的研究旨趣，进行贴切的时间编码设计（Biesanz，Deeb-Sossa，Papadakis，Bollen，Curran，2004）。另外，SEM 成长曲线模型亦可进行跨领域分析（Willett & Sayer，1996）及加入中介变量、调节变量与延伸到多变量模型（Curran & Hussong，2003），SEM 可谓多才多艺。

第七节 多重指标潜在特质成长模型

一、意义

前述的潜在特质成长模型,在每一测量的时间点上均只有一个测量指标,如果在每一测量的时间点上均有多个测量指标,一般的研究者可能会通过加总而成立一个新变量。不过,Bandalos(2002)指出此种题组分析法(item parceling)可能导致偏差的参数估计值,尤其当每一测量点上的多个测量指标违反单一向度(unidimensionality)假设时。比较好的做法是将每一测量点上的多个测量指标视为出自于潜在因子或建构,而针对这些一阶潜在因子进行成长模型分析。Hancock, Kuo, Lawrence(2001)称之为二阶成长模型(second-order growth modeling)。事实上此种模型有许多不同名称,例如,多重指标潜在特质成长模型(multiple indicator latent growth model, Chang, 1988)、潜在变量纵贯式曲线模型(latent variable longitudinal curve model, Tisak & Meredith, 1990)、曲线因子模型(factor-of-curves model, McArdle, 1988)。此种二阶测量模型有一基本假设:纵贯性测量不变性之假设(measurement invariance across time or longitudinalmeasurement invariance),意指这些多重指标在不同的时间测量点上,均在测同一特质。

二、理论基础

为具体说明多重指标潜在特质成长模型的基本内涵,兹以美国教育统计中心所赞助的 1988 年全国性纵贯性研究资料为例,利用 Amos 制作如图 5-35 的美国 791 位女学生自我概念的线性成长模型(引自 Hancock, Kuo, & Lawrence, 2001)。该研究旨在探究学生在 8 年级、10 年级、12 年级时自我概念的发展情形。该问卷含有三个有关自我概念的问题:

A. 整体而言,我感觉身心愉快。
D. 我觉得我是一个有价值的人。
H. 整体而言,我对自己感到满意。

图 5-35 中 44A,44D,44H 等即表示测量前述三个自我概念的观察指标。

由图 5-35 亦知,此二阶二因子 LCA 模型含有一阶的三个因素:SC8,SC10,SC12 及二阶的截距与成长因素:Level 与 Shape。利用图 5-35 中相关的参数代码,各指标的一阶测量回归模型:$Y = \tau + \Lambda\eta + \varepsilon$,可具体陈述如下:

44A = 0 + 1 * SC8 + E1。由此关系可推知 44A 指标的期望值会等于 SC8 的期望值;同理,其余指标的一阶测量回归模型为:

44D = I + d * SC8 + E2;
44H = J + h * SC8 + E3;
62A = 0 + 1 * SC10 + E4;
62D = I + d * SC10 + E5;
62H = J + h * SC10 + E6;
66A = 0 + 1 * SC12 + E7;

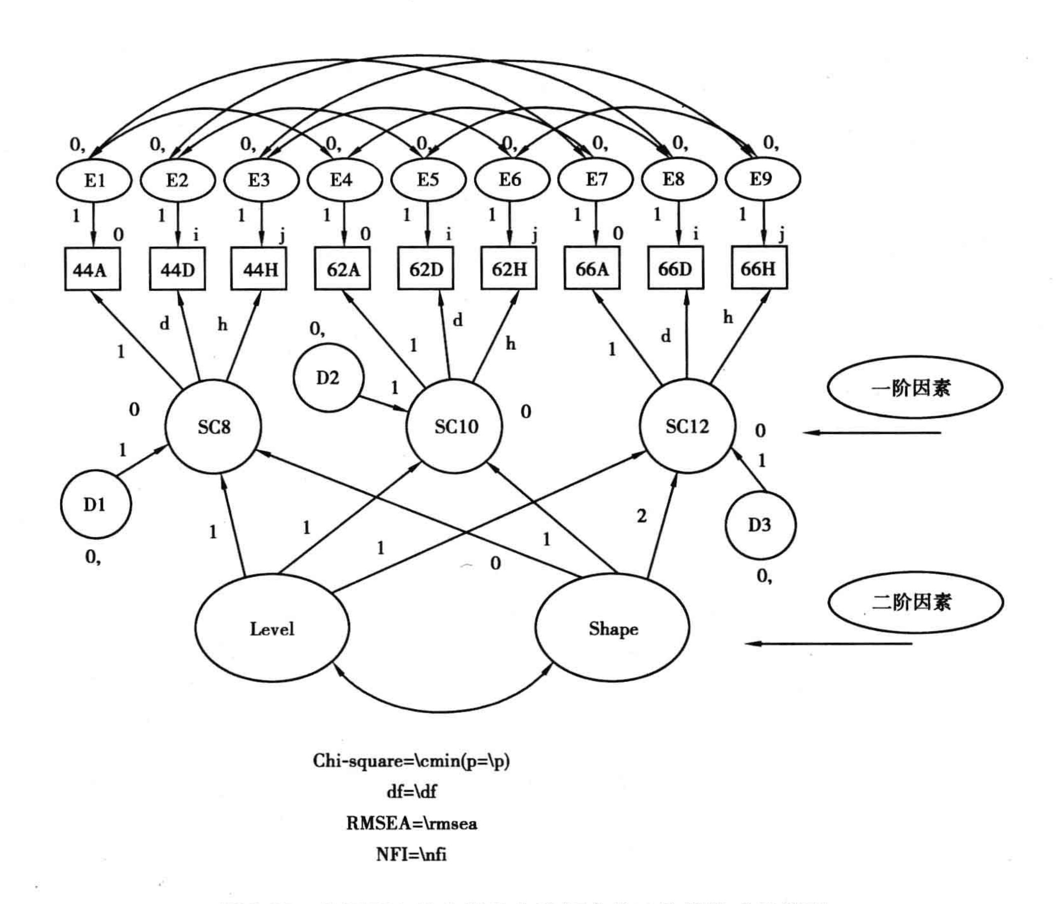


图 5-35 美国 791 位女学生自我概念的二阶线性成长模型

$$66D = I + d * SC12 + E8;$$
$$66H = J + h * SC12 + E9。$$

而其二阶测量之回归模型：

SC8 = 1 * LEVEL + 0 * SHAPE + D1。由此关系可推知 Level 的期望值会等于 SC8 或 44A 指标的期望值；同理，

$$SC10 = 1 * LEVEL + 1 * SHAPE + D2;$$
$$SC12 = 1 * LEVEL + 2 * SHAPE + D3。$$

利用上述公式，即可计算出指标变量与潜在变量的期望值。这些相关变量的 SEM 平均数结构的理论模型，透露出以下五项参数限制要点需要进行设定：

1. 该成长模型为线性，所以 Shape 的相关路径系数分别设定为 0,1,2。
2. 第一个测量指标设定为第一阶因素的量尺指标，其他的相对应指标的路径系数设定为相等（例如：d,h），以符合测量不变性的基本假设。
3. 在不同测量点上的相对应指标的误差设定为相关（例如：E1,E4,E7 间具有相关）。
4. Level 与 Shape 设定为具有相关。
5. 除了第 1,4,7 指标（其截距均设定为 0）之外，在不同测量点上的相对应指标的截距均设定为相等（例如：i,j）。

三、实例解说：单样本二阶多重指标潜在特质成长模型

以下兹以 Ingelet 等人(1994)所做的美国教育统中心赞助的 1988 年全国性纵贯性研究资料为例(引自 Hancock, Kuo, & Lawrence, 2001)，简介如何运用 Amos 分析多重指标潜在特质的成长模型。

该研究涉及 791 位女学生与 764 位男学生，这些学生被要求以四点量尺：1～4(非常不同意～非常不同意)，回答上述自我概念的问题，分数越高自我概念越强。该研究女学生的相关矩阵与描述统计等数据，摘要于图 5-36 的 Excel 表单中。

	A	B	C	D	E	F	G	H	I	J	K
1	rowtype_	varname_	44A	44D	44H	62A	62D	62H	66A	66D	66H
2	N		791	791	791	791	791	791	791	791	791
3	CORR	44A	1								
4	CORR	44D	0.46	1							
5	CORR	44H	0.523	0.458	1						
6	CORR	62A	0.201	0.147	0.173	1					
7	CORR	62D	0.147	0.143	0.152	0.818	1				
8	CORR	62H	0.134	0.132	0.163	0.835	0.831	1			
9	CORR	66A	0.095	0.055	0.14	0.119	0.123	0.137	1		
10	CORR	66D	0.105	0.091	0.141	0.102	0.131	0.131	0.866	1	
11	CORR	66H	0.076	0.051	0.138	0.098	0.125	0.141	0.888	0.839	1
12	MEAN		3.1	3.17	3	2.96	2.97	2.81	2.88	2.91	2.8
13	STDDEV		0.93	1.1	1.16	1.33	1.42	1.4	1.6	1.67	1.65

图 5-36 美国 791 位女学生自我概念的相关矩阵数据

进行单样本二阶潜在特质成长模型的分析，首先须在 Amos 之路径图设计窗口中，建构如同图 5-35 的理论模型，之后再利用 Amos 的 Data Files 窗口连接图 5-36 的数据文件之后，就可执行 SEM 统计分析，进而获得如图 5-37 的结果。

由图 5-37 底部的统计检验结果知，提议的线性成长模型与女学生的自我概念数据适配度极佳($\lambda^2=35.382$，$p=0.063$，RMSEA $=0.025$，NFI $=0.993$)。请注意截距因素 Level 的变异估计值 0.2555($z=3.04$，$p=0.0023$)，反映出这些女学生的自我概念，在初测时即存在着显著的个别差异；而初测之后自我概念的成长比率变异(0.0751)在个别间并无显著差异($p=0.1683$)。因为 Level 的期望值等于 44A 指标的期望值，所以图中 Level 的自我概念的起始值 3.10 即等于 44A 指标的样本平均数。这些女学生的自我概念，在初测之后呈现成长比率显著下降的现象(-0.1216，$z=-4.3882$，$p=0.000$)，读者从表 5-21 各指标的理论平均数很容易地看出来(从 3.1024 降到 2.7452)。

表 5-21 各指标在 8，10，12 年级的理论平均数

66H	66D	66A	62H	62D	62A	44H	44D	44A
2.7452	2.8918	2.8593	2.8704	3.0153	2.9808	2.9957	3.1387	3.1024

但这只就一般平均值来看，并不能说全部的女学生都是如此。如从 ShapeV 的方差(0.0751)观之，将之开平方根得到标准差为 0.274，此值比成长因素的平均数还大，反映出可能尚有一些女学生的自我概念，在初测之后呈现显著上升。

另外，由图 5-38 知，这些女学生自我概念的初测水平与之后自我概念的线性成长关

系,并不具有显著的相关(r = -0.35,z = -0.865,p =0.3870)。

表5-22 与表5-23 是图5-37 路径图中的一阶因子对各指标路径系数与截距的显著性检验结果,为符合纵贯式测量的不变性基本假设,请注意表中标签一栏(例如 d,h,i,j)的估计值均相同。

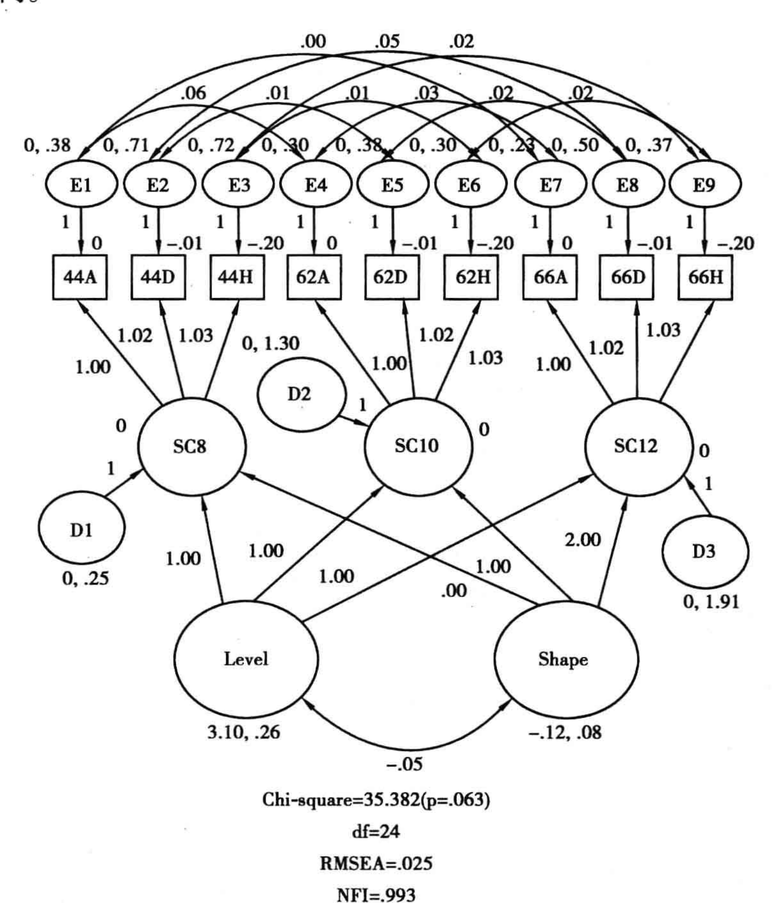


图 5-37 美国 791 学生自我概念的线性成长模型的 Amos 分析结果

表 5-22 一阶因子对各指标之回归系数摘要表

			Estimate	S. E.	C. R.	P	Label
44D	<---	SC8	1.0156	.0163	62.1999	***	d
44H	<---	SC8	1.0304	.0156	66.1868	***	h
62A	<---	SC10	1.0000				
62D	<---	SC10	1.0156	.0163	62.1999	***	d
62H	<---	SC10	1.0304	.0156	66.1868	***	h
66A	<---	SC12	1.0000				
66D	<---	SC12	1.0156	.0163	62.1999	***	d
66H	<---	SC12	1.0304	.0156	66.1868	***	h

表 5-23　各指标之截距摘要表

	Estimate	S. E.	C. R.	P	Label
44D	-.0121	.0522	-.2325	.8161	i
44H	-.2011	.0497	-4.0496	* * *	j
62D	-.0121	.0522	-.2325	.8161	i
62H	-.2011	.0497	-4.0496	* * *	j
66D	-.0121	.0522	-.2325	.8161	i
66H	-.2011	.0497	-4.0496	* * *	j

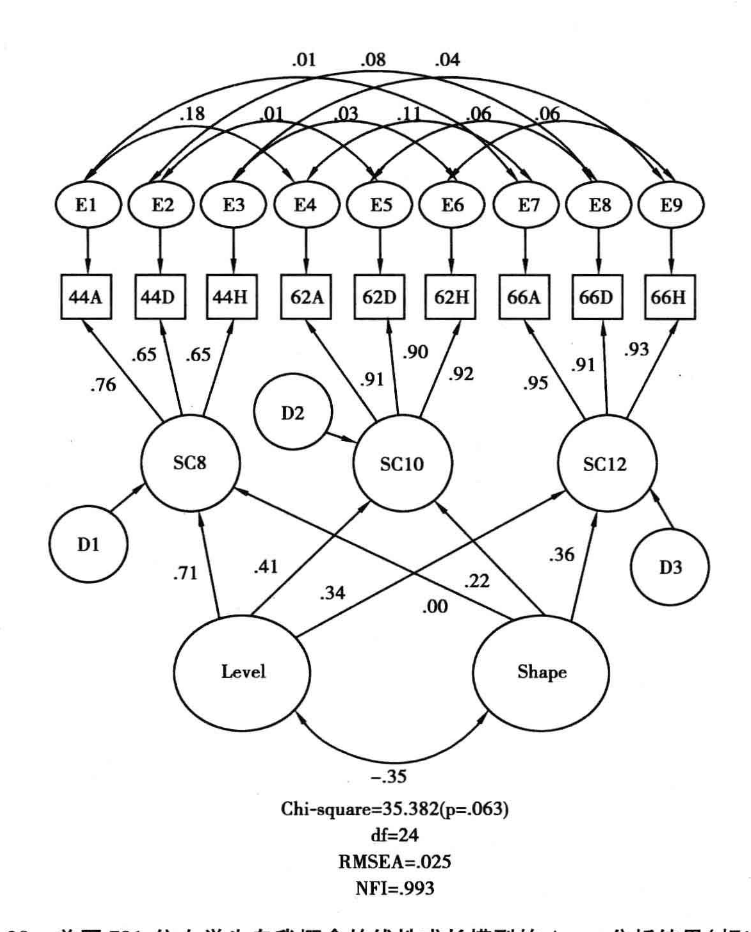


图 5-38　美国 791 位女学生自我概念的线性成长模型的 Amos 分析结果(标准化)

四、实例解说:多样本二阶多重指标潜在特质成长模型

多样本二阶多重指标潜在特质成长模型在组间的不变性检验,可以用来比较二阶成长因子的相关参数估计值。本实例仍沿用 Ingel 等人(1994)所做由美国教育统计中心赞助的 1988 年全国性纵贯性研究资料(引自 Hancock, Kuo, & Lawrence,2001),简介如何运用 Amos 分析双样本多重指标潜在特质的成长模型。该研究涉及 791 位女学生与 764 位男学生,这些学生被要求以四点量尺:1 ~4(非常不同意 ~ 非常不同意),回答自我概念之问题,分数越高自我概念越强。该研究男学生的相关矩阵与描述统计等数据,摘要于

图 5-39 的 Excel 表单中，该研究女学生的相关矩阵与描述统计等数据，则已摘要于图5-36 的 Excel 表单中。

	A	B	C	D	E	F	G	H	I	J	K
1	rowtype_	varname_	44A	44D	44H	62A	62D	62H	66A	66D	66H
2	N		764	764	764	764	764	764	764	764	764
3	CORR	44A	1								
4	CORR	44D	0.36	1							
5	CORR	44H	0.514	0.51	1						
6	CORR	62A	0.114	0.09	0.034	1					
7	CORR	62D	0.069	0.11	0.04	0.81	1				
8	CORR	62H	0.075	0.1	0.049	0.78	0.745	1			
9	CORR	66A	0.009	0.08	0.047	0.21	0.193	0.193	1		
10	CORR	66D	-0.01	0.07	0.048	0.2	0.213	0.206	0.925	1	
11	CORR	66H	0.014	0.09	0.078	0.2	0.198	0.203	0.938	0.936	1
12	MEAN		3.38	3.28	3.26	3.14	3.08	2.9	2.72	2.62	2.56
13	STDDEV		0.76	1.1	0.91	1.26	1.36	1.43	2.12	2.15	2.1

图 5-39　美国 791 位男学生自我概念的相关矩阵数据

以下简介多样本二阶潜在特质成长模型的 Amos 操作步骤：

(一) 绘制如图 5-40 之路径图

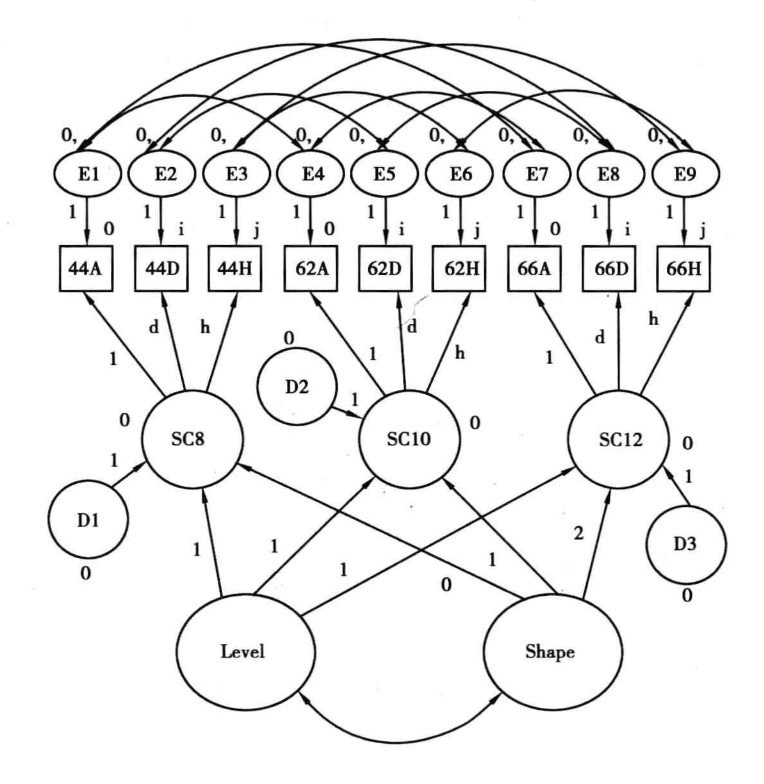


图 5-40　二阶潜在特质成长模型路径图

绘制此路径图须考虑以下的假设与限制：

1. 假定成长模型为线性，所以 Shape 的相关路径系数分别为 0，1，2。

2. 第一个测量指标设定为第一阶因素的量尺指标，其他的相对应指标的路径系数设定为相等（如 d，h 之相同标签），以符合测量不变性的基本假设。

3. 在不同测量点上的相对应指标的误差设定为相关（例如 E1，E4，E7 间具有相关）。

4. Level 与 Shape 设定为具有相关。

5. 除了第 1，4，7 指标（其截距均设定为 0）之外。在不同测量点上的相对应指标的截距设定为相等（如 i，j 之相同标签）。

（二）利用功能表单[Analyze]下的组别管理窗口[Manage Groups]建立男女生群组

建立不同群组时，使用者须在图 5-41 与图 5-42 中[Group Name]下界定群组名称。

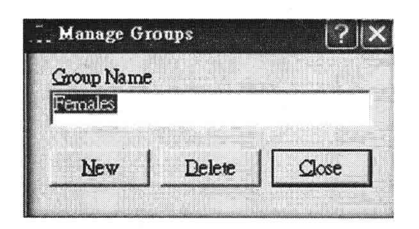


图 5-41　女生组

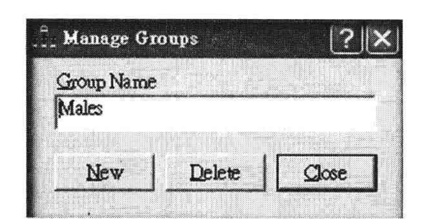


图 5-42　男生组

（三）利用功能表单[Analyze]下的多群组窗口[Multiple-Groups Analysis]（参见图 5-43），设定相关参数之标签。

因为此二阶潜在特质成长模型只涉及测量模型，因而只需在图 5-43 的窗口中，勾选[Measurement weighted]与[Measurement intercepts]两项，其他结构参数部分则不予勾选。按下[OK]之后，Amos 就会自动将待估计的参数加以命名，参见图 5-44 与图 5-45。

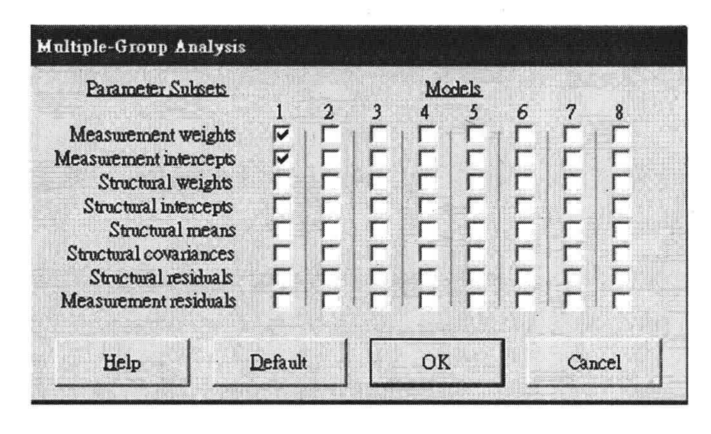


图 5-43　多群组窗口内估计参数的设定

（四）利用功能表单[File]下之[Data Files]连接双样本的数据文件

当上述图 5-36 与图 5-39 的数据文件建立后，按下 Amos[FILE]下之[DATA FILES]，

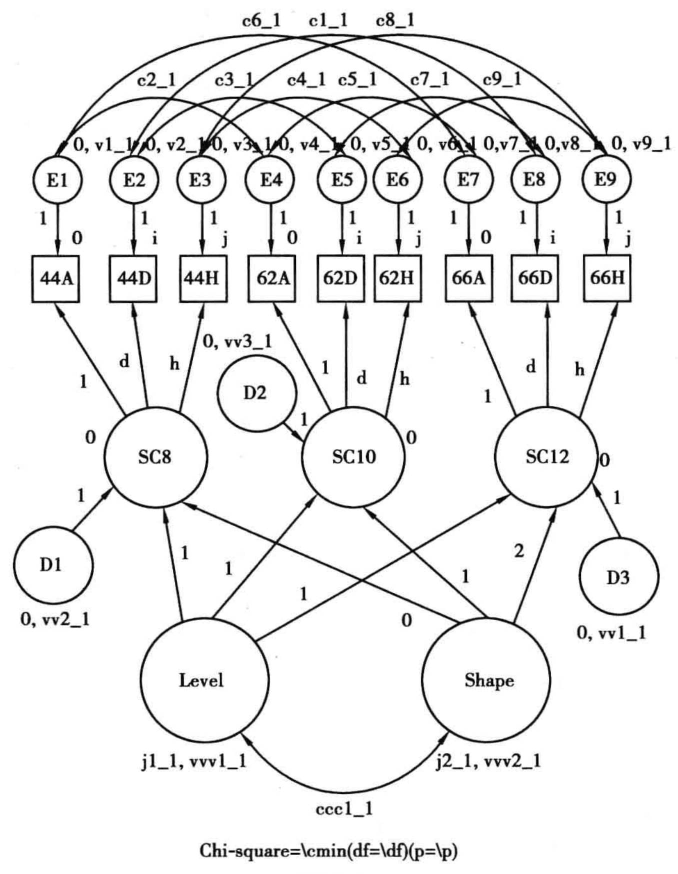


图 5-44　多样本二阶多重指标潜在特质成长模型：女生编码

出现图 5-46 之窗口后，点选[File Name]读入 Excel 所建的文件。当男、女生待分析的数据文件名称出现在此窗口之中，即表示 Amos 已可将路径图与此数据文件相互连接，进行统计分析。

（五）利用功能表单[View]下之[Analysis Properties]中，点选[Output]设定相关之输出统计量。

研究者如需要 Amos 输出标准化参数估计值与组别间参数之差异性检验值，需在图 5-47 分析属性窗口内勾选[Standardized estimates]与[Critical ratios fordifferences]。

（六）在[Analyze]下点选[Calculate Estimates]，执行统计分析与解释

请注意图 5-48 与表 5-36 的截距因素 Level 的变异估计值 0.2566（$z = 3.0546$，$p = 0.0023$），反映出女学生的自我概念在初测时即存在著显著的个别差异；而初测之后自我概念的成长比率变异（0.0759）在个别间并无显著差异（$z = 1.3892$，$p = 0.1648$）。从图 5-48 底部的显著性检验（$\chi^2 = 124.123$，df = 52，CFI = 0.993，RMSEA = 0.030）与适配度指标可推知，本模型的适配度尚佳。如依据 Steiger（1988）的校正建议，上式的 RMSEA =

0.042,须在多样本上 RMSEA 加以校正:将原本所获得的 RMSEA 乘以$\sqrt{k}$(k 代表组别数),亦即 $0.0299 * \sqrt{2} = 0.042$。

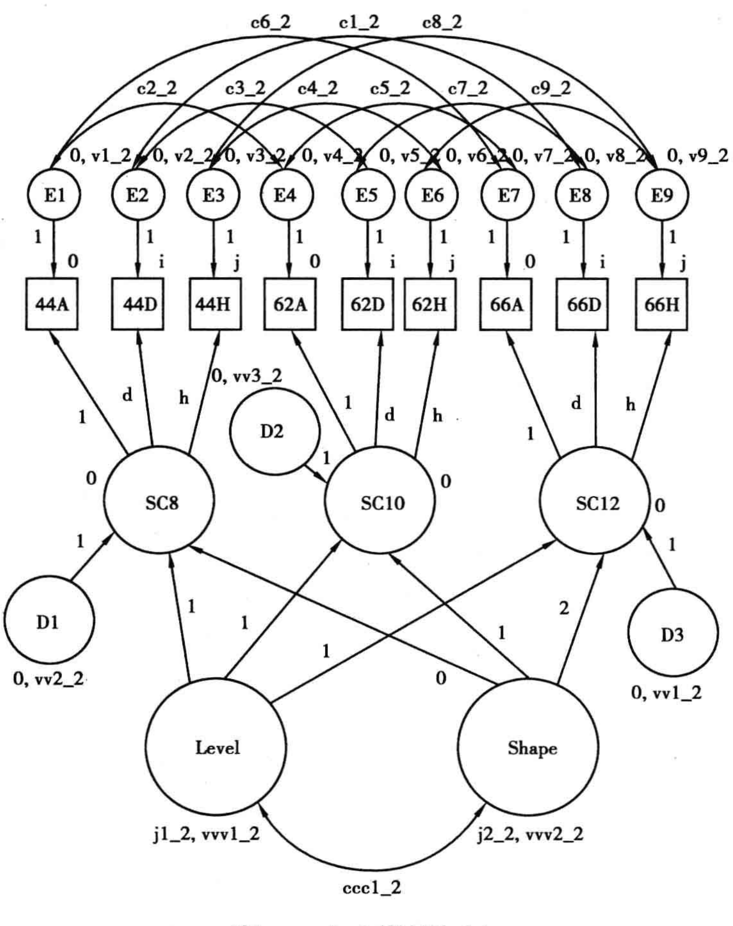


图 5-45　多样本二阶多重指标潜在特质成长模型:男生编码

Data Files

Group Name	File	Variable	Value	N
Females	FEMALE (XLS)			791/791
Males	MALE (XLS)			764/764

File Name　Working File　Help
View Data　Grouping Variable　Group Value
OK　Cancel

图 5-46　男女生数据文件之连接

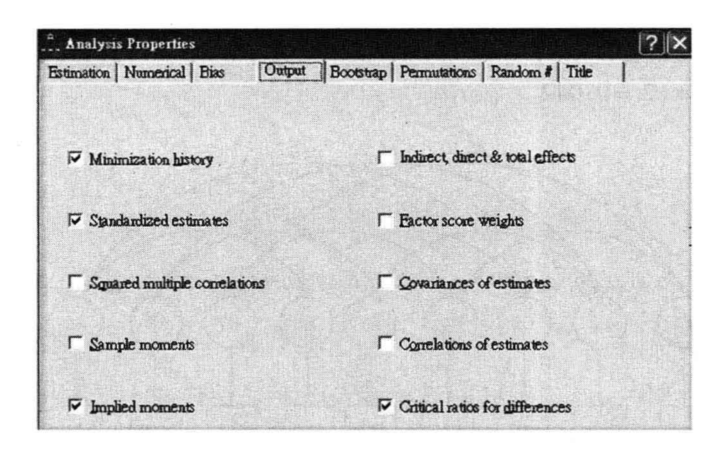


图 5-47 分析属性窗口内所需统计量的输出设定

由图 5-49 与表 5-35 的统计检验结果知，男生的截距因素 Level 的变异估计值0.0794 ($z = 1.1114, p = 0.2664$)，反映出这些男学生的自我概念，在初测时个别间的变异就不大；而初测之后自我概念的成长比率变异(0.2364, $z = 3.6202, p = 0.000$)，在个别间具有显著差异。

表 5-24 男女生在截距与斜率上的平均数与显著性检验

		Estimate	S. E.	C. R.	P	Label
女	Level	3.1268	.0298	104.9911	* * *	j1_1
生	Shape	-.1216	.0277	-4.3849	* * *	j2_1
男	Level	3.3729	.0255	132.0156	* * *	j1_2
生	Shape	-.3105	.0342	-9.0666	* * *	j2_2

表 5-24 的数据，是针对各主要参数间之差异性的检验(Critical Ratios for Differences Between Parameters, CR 值检验，请利用[分析属性窗口]中的 Output 窗口中进行设定)。因为 CR 值部分报表很长，不易在一页中完全呈现，仅摘述重要结果与评论如下：

1. CR(J1 - 1 vs J1 - 2) = 6.44，反映出男学生的截距因子的平均数 3.373 显著高于女学生的截距因子的平均数 3.127 ($p < 0.05$)，两者差异的效果值为

$$0.597\left(=\frac{|3.3729-3.1268|}{\sqrt{\dfrac{791*0.257+764*.0.79}{791+764}}}\text{，依据 Cohen(1988)之定义，约为中效果值}\right),$$

亦即一般而言男学生在 8 年级时之自我概念比女学生约高出半个标准差。

2. CR(J2 - 1 vs J2 - 2) = -4.29，反映出男学生的成长因子的平均数 -0.310 显著高于女学生的成长因子的平均数 -0.122($p<0.05$)，两者差异的效果值为0.478

$$\left(=\frac{|-0.3105-(-0.1216)|}{\sqrt{\dfrac{791*0.076+764*0.236}{791+764}}}\text{，依据 Cohen(1988)的定义，约为中效果值}\right)\text{，亦即}$$

一般而言，男学生的自我概念在 10 与 12 年级时，比女学生的自我概念的下降比率快约半个标准差。

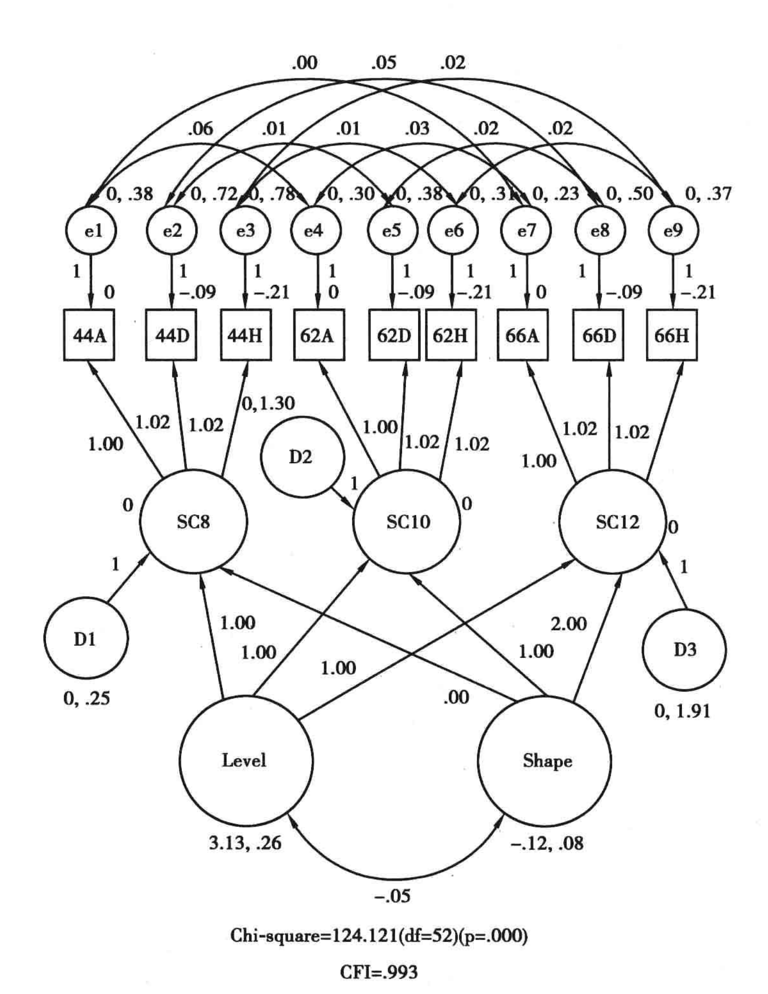


图5-48　多样本二阶多重指标潜在特质成长模型分析结果:女生

至于截距因子与成长因子间在男女组别间的相关差异性[CR(CCC1 -1vs CCC1 -2) =0.57]及变异之差异性[CR(VVV1 -1 vs VVV1 -2) = -1.61,CR(VVV2 -1 vs VVV2 -2) = 1.89]均未达0.05 的显著水平。

由表5-25 的结果知,这些男学生的自我概念,在初测之后呈现成长比率显著下降的现象(-0.3105, z = -9.07,p = 0.000),读者从表5-25 各指标的理论平均数就能轻易地看出来。但这只就一般平均值来看,并不能说全部的男学生都是如此。如从 Shape 的方差(0.2364)观之,将之开平方根得到标准差为0.486,反映出可能尚有一些男学生的自我概念,在初测之后呈现显著上升。

由表5-26 的结果知,这些女学生的自我概念,在初测之后呈现成长比率显著下降的现象(-0.1216,z = -4.3882,p =0.000),读者从表5-26 各指标的理论平均数就能看出来。但这只就一般平均值来看,并不能说全部的女学生都是如此。如从 Shape 的方差(0.0759)观之,将之开平方根得到标准差为0.275,此值比成长因素的平均数还大,反映出可能尚有一些女学生的自我概念,在初测之后呈现显著上升。

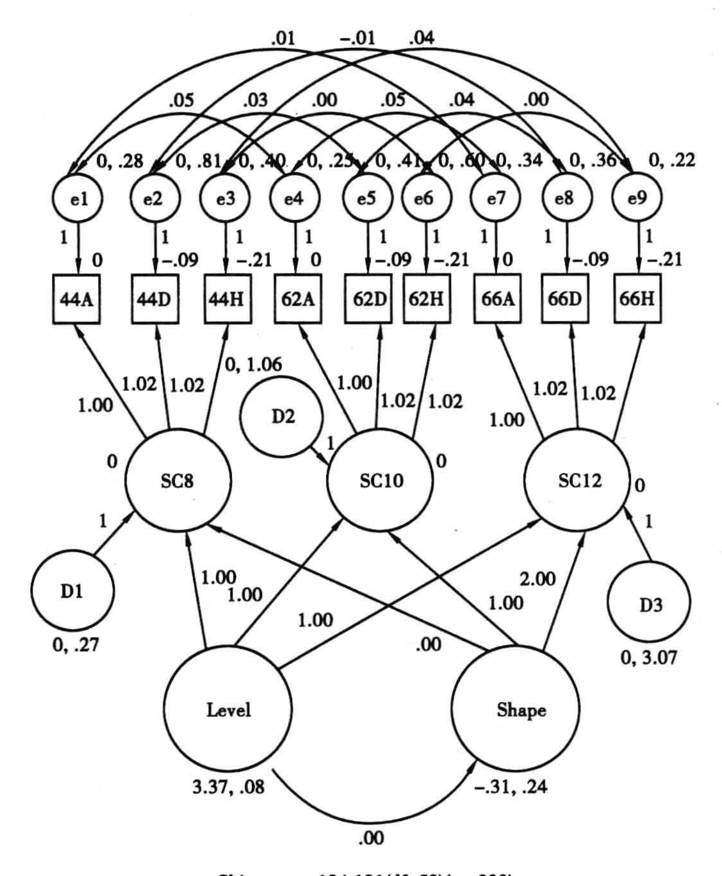


图 5-49 多样本二阶多重指标潜在特质成长模型分析结果:男生

表 5-25 男学生的自我概念在各阶段多重指标的理论平均数

66H	66D	66A	62H	62D	62A	44H	44D	44A
2.6083	2.7220	2.7520	2.9257	3.0391	3.0624	3.2432	3.3562	3.3729

表 5-26 女学生的自我概念在各阶段多重指标的理论平均数

66H	66D	66A	62H	62D	62A	44H	44D	44A
2.7429	2.8565	2.8836	2.8672	2.9807	3.0052	2.9916	3.1049	3.1268

为让读者更清楚地掌握住前述路径图中的相关参数估计值及其相关的统计检验结果,兹将 Amos 的统计输出结果呈现在表 5-27 ~ 表 5-36 中。请注意表中各参数的标签(label)如果相同(例如 d, h, i, j),亦即这些参数进行了等同限制。因此,这些相关的参数估计值均相等。

表 5-27　Regression Weights：(Males-Unconstrained)

			Estimate	S. E.	C. R.	P	Label
44D	<---	SC8	1.0213	.0104	98.1990	* * *	d
44H	<---	SC8	1.0224	.0097	105.1256	* * *	h
62A	<---	SC10	1.0000				
62D	<---	SC10	1.0213	.0104	98.1990	* * *	d
62H	<---	SC10	1.0224	.0097	105.1256	* * *	h
66A	<---	SC12	1.0000				
66D	<---	SC12	1.0213	.0104	98.1990	* * *	d
66H	<---	SC12	1.0224	.0097	105.1256	* * *	h

表 5-28　Regression Weights (Females-Unconstrained)

			Estimate	S. E.	C. R.	P	Label
44D	<---	SC8	1.0213	.0104	98.1990	* * *	d
44H	<---	SC8	1.0224	.0097	105.1256	* * *	h
62A	<---	SC10	1.0000				
62D	<---	SC10	1.0213	.0104	98.1990	* * *	d
62H	<---	SC10	1.0224	.0097	105.1256	* * *	h
66A	<---	SC12	1.0000				
66D	<---	SC12	1.0213	.0104	98.1990	* * *	d
66H	<---	SC12	1.0224	.0097	105.1256	* * *	h

表 5-29　Intercepts：(Males-Unconstrained)

	Estimate	S. E.	C. R.	P	Label
44D	-.0885	.0343	-2.5829	.0098	i
44H	-.2052	.0321	-6.4033	* * *	j
62D	-.0885	.0343	-2.5829	.0098	i
62H	-.2052	.0321	-6.4033	* * *	j
66D	-.0885	.0343	-2.5829	.0098	i
66H	-.2052	.0321	-6.4033	* * *	j

表 5-30 Intercepts:(Females-Unconstrained)

	Estimate	S. E.	C. R.	P	Label
44D	-.0885	.0343	-2.5829	.0098	i
44H	-.2052	.0321	-6.4033	***	j
62D	-.0885	.0343	-2.5829	.0098	i
62H	-.2052	.0321	-6.4033	***	j
66D	-.0885	.0343	-2.5829	.0098	i
66H	-.2052	.0321	-6.4033	***	j

表 5-31 Covariances:(Males)

			Estimate	S. E.	C. R.	P	Label
Level	<-->	Shape	-.0041	.0540	-.0763	.9392	cccl_2
E2	<-->	E8	-.0065	.0246	-.2631	.7925	c1_2
E1	<-->	E4	.0465	.0165	2.8241	.0047	c2_2
E2	<-->	E5	.0252	.0267	.9455	.3444	c3_2
E3	<-->	E6	.0004	.0232	.0166	.9868	c4_2
E4	<-->	E7	.0495	.0173	2.8660	.0042	c5_2
E1	<-->	E7	.0094	.0166	.5648	.5722	c6_2
E5	<-->	E8	.0435	.0198	2.1951	.0282	c7_2
E3	<-->	E9	.0396	.0171	2.3179	.0205	c8_2
E6	<-->	E9	-.0016	.0201	-.0783	.9376	c9_2

表 5-32 Covariances:(Females)

			Estimate	S. E.	C. R.	P	Label
Level	<-->	Shape	-.05	.06	-.87	.38	cccl_1
E2	<-->	E8	.05	.03	1.94	.05	c1_1
E1	<-->	E4	.06	.02	3.28	.00	c2_1
E2	<-->	E5	.01	.02	.37	.71	c3_1
E3	<-->	E6	.01	.02	.62	.53	c4_1
E4	<-->	E7	.03	.02	1.92	.06	c5_1
E1	<-->	E7	.00	.02	.25	.80	c6_1
E5	<-->	E8	.02	.02	1.18	.24	c7_1
E3	<-->	E9	.02	.02	.80	.42	c8_1
E6	<-->	E9	.02	.02	1.02	.31	c9_1

表 5-33　Correlations：(Males)

			Estimate
Level	<-->	Shape	-.0301
E2	<-->	E8	-.0120
E1	<-->	E4	.1764
E2	<-->	E5	.0435
E3	<-->	E6	.0008
E4	<-->	E7	.1721
E1	<-->	E7	.0305
E5	<-->	E8	.1125
E3	<-->	E9	.1348
E6	<-->	E9	-.0044

表 5-34　Correlations：(Females)

			Estimate
Level	<-->	Shape	-.3476
E2	<-->	E8	.0882
E1	<-->	E4	.1829
E2	<-->	E5	.0175
E3	<-->	E6	.0308
E4	<-->	E7	.1190
E1	<-->	E7	.0156
E5	<-->	E8	.0571
E3	<-->	E9	.0384
E6	<-->	E9	.0557

表 5-35　Variances：(Males)

	Estimate	S. E.	C. R.	P	Label
Level	.0794	.0715	1.1114	.2664	vvv1_2
Shape	.2364	.0653	3.6202	***	vvv2_2
D3	3.0692	.2431	12.6260	***	vv1_2
D1	.2729	.0742	3.6765	***	vv2_2
D2	1.0566	.0766	13.7938	***	vv3_2
E1	.2828	.0229	12.3479	***	v1_2
E2	.8090	.0476	17.0100	***	v2_2
E3	.3978	.0280	14.1995	***	v3_2
E4	.2463	.0247	9.9707	***	v4_2
E5	.4144	.0310	13.3609	***	v5_2
E6	.5972	.0387	15.4278	***	v6_2
E7	.3363	.0249	13.4891	***	v7_2
E8	.3611	.0264	13.6786	***	v8_2
E9	.2171	.0219	9.9228	***	v9_2

表 5-36 Variances: (Females)

	Estimate	S. E.	C. R.	P	Label
Level	.2566	.0840	3.0546	.0023	vvl_1
Shape	.0759	.0546	1.3892	.1648	vv2_1
D3	1.9131	.1722	11.1068	* * *	vl_1
D1	.2533	.0856	2.9582	.0031	v2_1
D2	1.30	.0829	15.7370	* * *	v3_1
E1	.38	.0303	12.5047	* * *	v1_1
E2	.72	.0452	15.8634	* * *	v2_1
E3	.7277	.0457	15.9246	* * *	v3_1
E4	.3017	.0238	12.6904	* * *	v4_1
E5	.3818	.0273	14.0098	* * *	v5_1
E6	.3092	.0246	12.5603	* * *	v6_1
E7	.2314	.0232	9.9870	* * *	v7_1
E8	.4973	.0330	15.0840	* * *	v8_1
E9	.3723	.0283	13.1766	* * *	v9_1

第八节 Amos Basic 程序

一、双因子潜在特质变化分析

```
#Region "Header"
Imports System
Imports System. Diagnostics
Imports Microsoft. VisualBasic
Imports AmosEngineLib
Imports AmosGraphics
Imports AmosEngineLib. AmosEngine. TMatrixID
Imports PBayes
#End Region
Module MainModule
        Sub Main( )
        Dim Sem As New AmosEngine
        Try
Sem. TableOutput
Sem. Standardized
Sem. ImpliedMoments( )
Sem. SampleMoments( )
```

```
Sem. ModelMeansAndIntercepts( )
Sem. BeginGroup( AmosEngine. AmosDir &"Examples\Lca1. xls" , "sheet1" )
Sem. AStructure( "y1  =  (1)Level  +  (0)Shape  +  (1)E1" )
Sem. AStructure( "y2  =  (1)Level  +  Shape  +  (1)E2" )
Sem. AStructure( "y3  =  (1)Level +  Shape  +  (1)E3" )
Sem. AStructure( "y4  =  (1)Level  +  (1)Shape  +  (1)E4" )
Sem. AStructure( "Level = (M_Int)  +  (1)D1" )
Sem. AStructure( "Shape = (M_Slp)  +  (1)D2" )
Sem. AStructure( "D1  <--> D2" )
Sem. FitModel
Finally
SEM. Dispose( )
End Try
End Sub
        End Module
```

二、双因子潜在特质变化之协方差分析

```
#Region "Header"
Imports System
Imports System. Diagnostics
Imports Microsoft. VisualBasic
Imports AmosEngineLib
Imports AmosGraphics
Imports AmosEngineLib. AmosEngine. TMatrixID
Imports PBayes
#End Region
Module MainModule
        Sub Main( )
        Dim Sem As New AmosEngine
        Try
Sem. TableOutput
Sem. Standardized
Sem. ImpliedMoments( )
Sem. SampleMoments( )
Sem. ModelMeansAndIntercepts( )
Sem. BeginGroup( AmosEngine. AmosDir &"Examples\Lca1. xls" , "sheet2" )
Sem. AStructure( "Y1 = (1)Level  +  (0)Shape  +  (1)E1" )
Sem. AStructure( "Y2 = (1)Level  +  (1)Shape  +  (1)E2" )
Sem. AStructure( "Y3 = (1)Level  +  (2)Shape  +  (1)E3" )
Sem. AStructure( "Y4 = (1)Level  +  (3)Shape  +  (1)E4" )
Sem. AStructure( "GENDER = (M_GENDER)  +  (1)E5" )
Sem. AStructure( "AGE = (M_AGE)  +  (1)E6" )
Sem. AStructure( "Level = (M_Int)  +  GENDER  +  AGE +  (1)D1" )
Sem. AStructure( "Shape = (M_Slp)  +  GENDER  +  AGE  +  (1)D2" )
```

```
Sem. AStructure(" D1  <-- >  D2")
Sem. AStructure(" E5  <-- >  E6")
Sem. FitModel
Finally
SEM. Dispose( )
End Try
End Sub
        End Module
```

本章习题

I. 概念题

1. Curran(1997)为了探究儿童反社会行为的发展对日后学业表现、犯罪、饮酒、过动等的影响,并欲了解儿童性别及家庭文化刺激与支持的预测效能,乃进行一纵贯式儿童反社会行为的研究,他提出图 5-50 的理论模型并获得相关回归系数的统计检验结果:如图 5-50 及表 5-37 所示。请根据此结果回答下列问题:

(1)Curran 所提出的理论模型与数据相适配吗?

(2)男、女童的反社会行为在第一次测量时具有显著的差异吗?

(3)男、女童的反社会行为在其后几次测量时的改变比率具有显著的差异吗?

(4)初期具有较高家庭文化刺激与支持的儿童,日后其反社会行为呈现递减还是递增的现象?

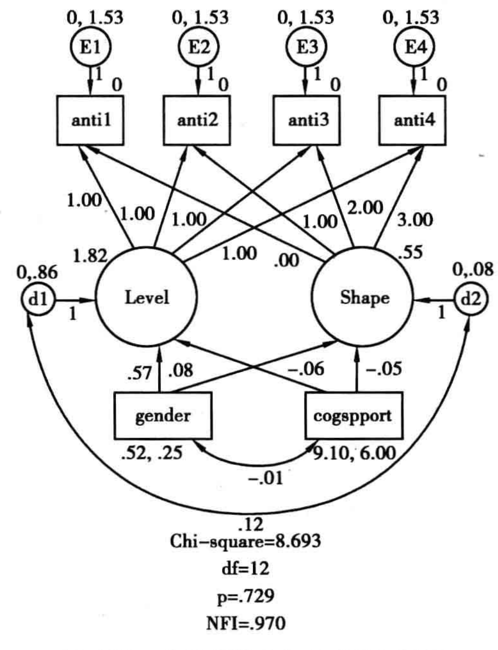


图 5-50 儿童反社会行为的发展模型

资料来源:Data taken from http://www.duke.edu/~curran

表 5-37　Regression Weights

			Estimate	S. E.	C. R.	P
Shape	<---	gender	.0791	.0841	.9403	.3470
Level	<---	cogspport	-.0626	.0382	-1.6378	.1015
Shape	<---	cogspport	-.0452	.0172	-2.6370	.0084
Level	<---	gender	.5716	.1875	3.0486	.0023

表 5-38　Implied Means

cogspport	gender	anti4	anti3	anti2	anti1
9.0995	.5249	2.0837	1.9072	1.7308	1.5543

表 5-39　Intercepts

	Estimate	S. E.	C. R.	P
Level	1.8240	.3740	4.8773	* * *
Shape	.5465	.1678	3.2573	.0011

(5)根据上述表 5-37 至表 5-39 Amos 的相关报表,计算这四次测量结果的期望值。

提示:

Mean(Anti1) = [Mean(Level) + Mean(gender) * (0.5716) + Mean(Cogspport) * (-0.0626)] + [0 * (Mean(Shape) + Mean(Gender) * (0.0791) + Mean(Cogspport) * (-0.0452))] + Mean(E1)

II. 应用题

1. 表 5-40 之体重相关矩阵,系小孩子分别在 5,7,9,11,13 岁时与母亲怀孕前体重的

表 5-40　小孩体重(3~13 岁)与母亲怀孕前体重的相关系数矩阵(N=155)

体重	1	2	3	4	5	6
1. Child at age 5	—					
2. Child at age 7	.7947	—				
3. Child at age 9	.7264	.8569	—			
4. Child at age 11	.6405	.7866	.8651	—		
5. Child at age 13	.6025	.7447	.7968	.8981	—	
6. Mother	.1592	.2891	.3550	.4175	.4296	—
M	39.5480	55.3160	72.3350	96.2520	119.1030	127.1670
SD	6.1096	11.1546	17.8567	26.9084	33.4412	21.2030

资料来源:Weight is given in pounds (1 lb. =0.45 kg). Baler, Keck, Mott, & Qoinlan (1993).

相关系数与平均数。为了取得小孩子在5岁时体重的估计值，研究者必须把时间的原点放在5岁。因此，时间编码时要将各阶段年龄减掉5岁。假设该理论模型为线性时，为了保留原来测量体重的两年间距，相关之 Lumbda 矩阵设计如下：

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 4 \\ 1 & 6 \\ 1 & 8 \end{bmatrix}$$

(1)请检验5~9岁间，孩子的体重成长是否呈现直线成长关系？

(2)请检验5~13岁间，孩子的体重成长是否呈现二次曲线关系？

提示：二次曲线关系的路径及系数设计，如图5-51的设计。

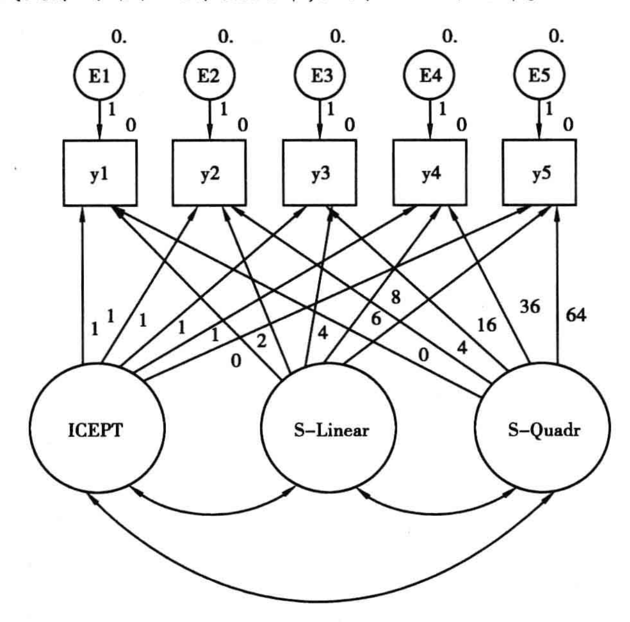


图5-51　二次曲线的体重成长模式

(3)假如研究者对于9岁时小孩体重的估计值较感兴趣，研究者就必须把时间的原点放在9岁。因此，时间编码时要将各阶段年龄减掉9岁。假设该理论模型为线性时，为了保留原来测量体重的两年间距，相关的 Lumbda 矩阵重新设计如下：

$$A = \begin{bmatrix} 1 & -4 \\ 1 & -2 \\ 1 & 0 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}$$

请检验此理论模型，并与(1)的检验结果作一比较。

2. 下图5-52、图5-54、图5-56数据(取自 Willett & Sayer, 1996)来自 NCDS(National Child Development Study)的纵贯式跨领域研究(Cross-domain analysis of change over time)，该研究针对英国儿童在出生时，被诊断为健康儿童(n = 514)及带有气喘(n = 437)与心脏病儿童(n = 72)，之后他们在7岁(入学年度)、11岁、16时，在阅读

(Reading)与数学(mathematics)上的线性成长的数据,分别建文件如图5-52—图5-56。

请根据以下三个数据:图5-53、图5-55、图5-57的分析结果,回答下列问题:

	A	B	C	D	E	F	G	H
1	rowtype_	varname_	R7	R11	R16	M7	M11	M16
2	mean		56.97	54.28	53.78	56.72	54.87	55.01
3	cov	R7	972.18					
4	cov	R11	583.33	823.82				
5	cov	R16	519.94	683.15	824.68			
6	cov	M7	431.75	369.76	371.11	773.68		
7	cov	M11	542.41	564.72	574.97	430.86	775.3	
8	cov	M16	448.83	504.99	549.05	380.74	563.86	764.4
9	n		514	514	514	514	514	514

图5-52　健康儿童阅读与数学的协方差矩阵

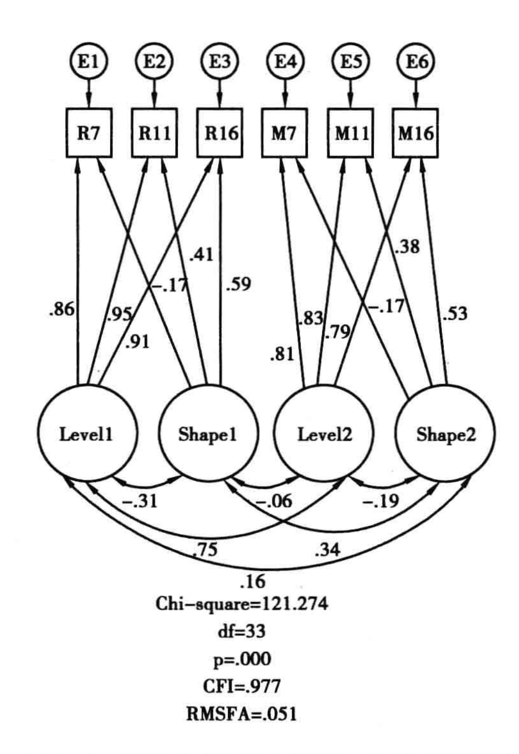


图5-53　健康儿童阅读与数学之线性成长模型

(1)Level1与Level2间具有高度相关,其涵意为何?

(2)为何学童在入学的第一年时,不管是阅读或是数学较好的学童,他们在其后的学业表现却呈现较差的成长速率?

(3)Shape1与Shape2间具有中度相关,其涵意为何?

(4)这个线性成长模型的假设成立吗?为什么?

(5)验证一下三组学童在阅读或是数学的期望值是否正确(参照表5-41—表5-43)。

(6)请检验一下这三组儿童的阅读与数学成长模型是否相同(Amos多群组分析)。

嘉義大學蚵教所, Fred Li, 2007 - Lca-cross-domain.xls

	A	B	C	D	E	F	G	H
1	rowtype_	varname_	R7	R11	R16	M7	M11	M16
2	mean		55.12	56.47	57.15	58.58	54.08	53.63
3	cov	R7	986.71					
4	cov	R11	548.25	802.96				
5	cov	R16	520.82	657.57	849.52			
6	cov	M7	481.1	372.74	381.45	800.32		
7	cov	M11	566.74	567.08	531	442.87	788.15	
8	cov	M16	463.27	494.04	546.13	349.44	589.6	819.31
9	n		437	437	437	437	437	437

图 5-54　气喘儿童阅读与数学的协方差矩阵

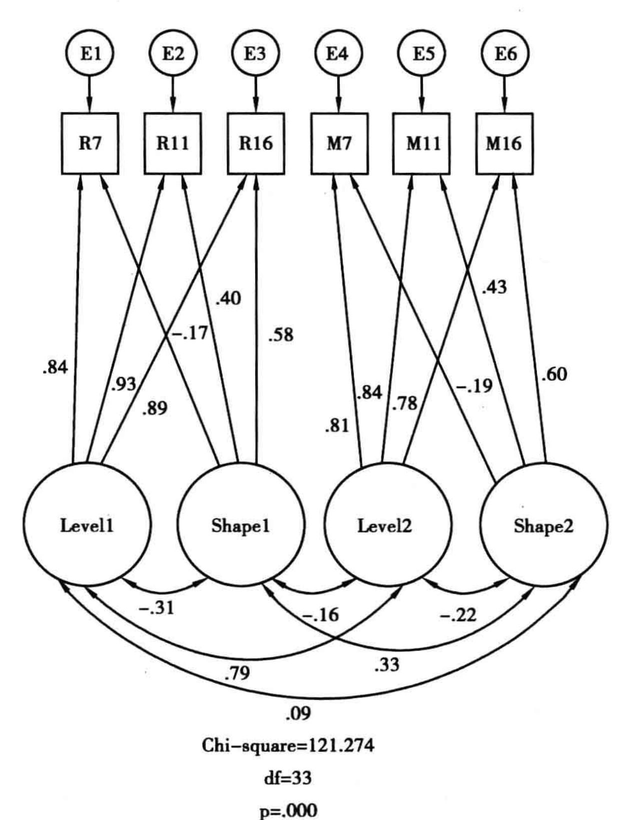


图 5-55　气喘儿童阅读与数学的线性成长模型

表 5-41　Implied Means（Healthy-Default model）

M16	M11	M7	R16	R11	R7
54.7458	55.2240	56.6302	53.6375	54.4709	56.9216

表 5-42 Implied Means(Asthma-Default model)

M16	M11	M7	R16	R11	R7
53.2589	54.5772	58.4538	57.0741	56.5717	55.0942

	A	B	C	D	E	F	G	H
1	rowtype_	varname_	R7	R11	R16	M7	M11	M16
2	mean		53.05	43.81	44.06	49.89	45.38	46.72
3	cov	R7	1155.83					
4	cov	R11	559.29	819.39				
5	cov	R16	573.27	709.01	943.51			
6	cov	M7	535.72	287.82	299.72	809.11		
7	cov	M11	608.91	580.76	541.5	421.66	773.57	
8	cov	M16	478.83	460.77	555.56	358.34	559.24	743.41
9	n		72	72	72	72	72	72

图 5-56 心脏病儿童阅读与数学之协方差矩阵

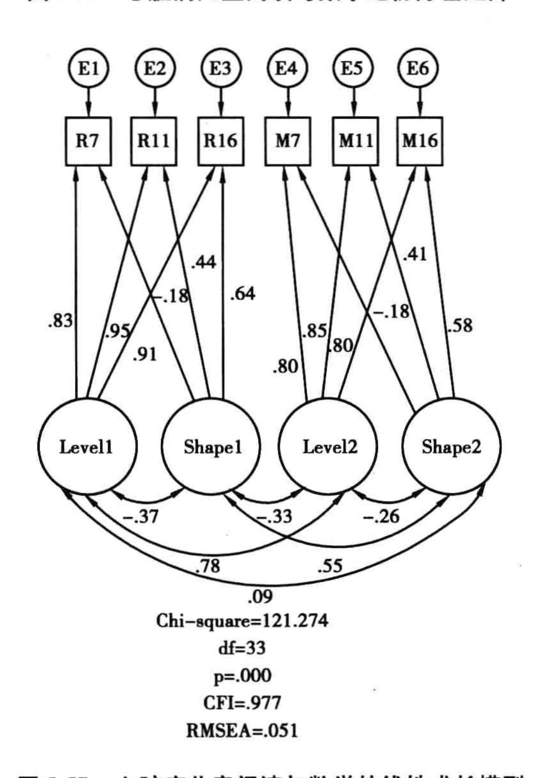


图 5-57 心脏病儿童阅读与数学的线性成长模型

表 5-43 Implied Means(Seizures-Default model)

M16	M11	M7	R16	R11	R7
45.7330	46.7027	49.5543	42.8949	45.3714	52.6538

3. 以下表 5-44 是荷兰 679 位学童的学业成就动机(含 x 与 y 两个指标)在 4 个不同时间所测得的资料,请回答下列问题:

(1)画出二阶多重指标线性成长模型路径图并进行适配度检验。

(2)进行纵贯性测量不等性(longitudinal measurement invariance)检验。

(3)解释分析结果。

表 5-44　Sample means snd covariance matrix of academic motivation N = 679

	Language							
	x_1	y_1	x_2	y_2	x_3	y_3	x_4	y_4
	3.604	3.640	3.518	3.248	3.443	3.209	3.446	3.241
x_1	.881							
y_1	.531	.763						
x_2	.213	.228	.850					
y_2	.260	.282	.551	1.069				
x_3	.166	.239	.268	.337	1.059			
y_3	.259	.315	.359	.453	.631	1.106		
x_4	.204	.231	.262	.336	.403	.427	1.121	
y_4	.216	.275	.287	.399	.367	.514	.720	1.101

资料来源:Driessen, Van Langen, & Vierke, 2000.

第六章 Amos Bayesian SEM 与 Bootstrapping 的应用及相对效能分析

本章旨在利用经验数据，具体说明 Bayesian SEM 与 Bootstrapping 在 Amos 中的操作过程，并分析两者在不同样本大小上的相对效能。兹以吴丽华(2007)的硕士论文为例，她利用过去的理论与文献，建立国小级任教师内外控信念、社会支持、教师工作压力与身心健康的关系模型如图6-1。图中所含的路径系数是由传统的SEM分析所计算出来的未标准化系数。

该模型内含有四个测量模型，各个测量模型的基本设定说明如下：

一、外因潜在变量“内外控信念”

外因潜在变量“内外控信念”是由二个外因观察变量来测量，包含内控与外控两个分量表，为解决“内外控信念”量尺未确定性的问题，乃将“内外控信念”潜在变量的方差设为1，以便估计各个观察变量的参数估计值。

二、外因潜在变量“社会支持”

外因潜在变量“社会支持”由四个外因观察变量来测量，包含同事支持、主任支持、家长支持与亲友支持等分量表，为解决“社会支持”量尺未确定性的问题，乃将“社会支持”潜在变量的方差设为1，以便估计各个观察变量的参数估计值。

三、中介潜在变量“教师工作压力”

中介潜在变量“教师工作压力”是由三个内因观察变量来测量，包含有工作负荷、专业知能与变革适应等分量表，为解决“教师工作压力”量尺未确定性的问题，乃将工作负荷指标的因素负荷量设为1。

四、内因潜在变量“身心健康”

内因潜在变量“身心健康”由三个内因观察变量来测量，包含生理症状、心理症状与社会功能不良等分量表，为解决“身心健康”量尺未确定性的问题，乃将“心理症状”指标的因素负荷量设为1。

以下将根据此理论模型与实际搜集的资料，简介如何利用 Amos 的 Bayesian SEM 进行统计分析。此外，并比较在不同样本大小上，Bayesian SEM 与 Bootstrapping 两种参数估计法的效能，以评估 Bayesian SEM 与 Bootstrapping 在不同样本上的相对效能。

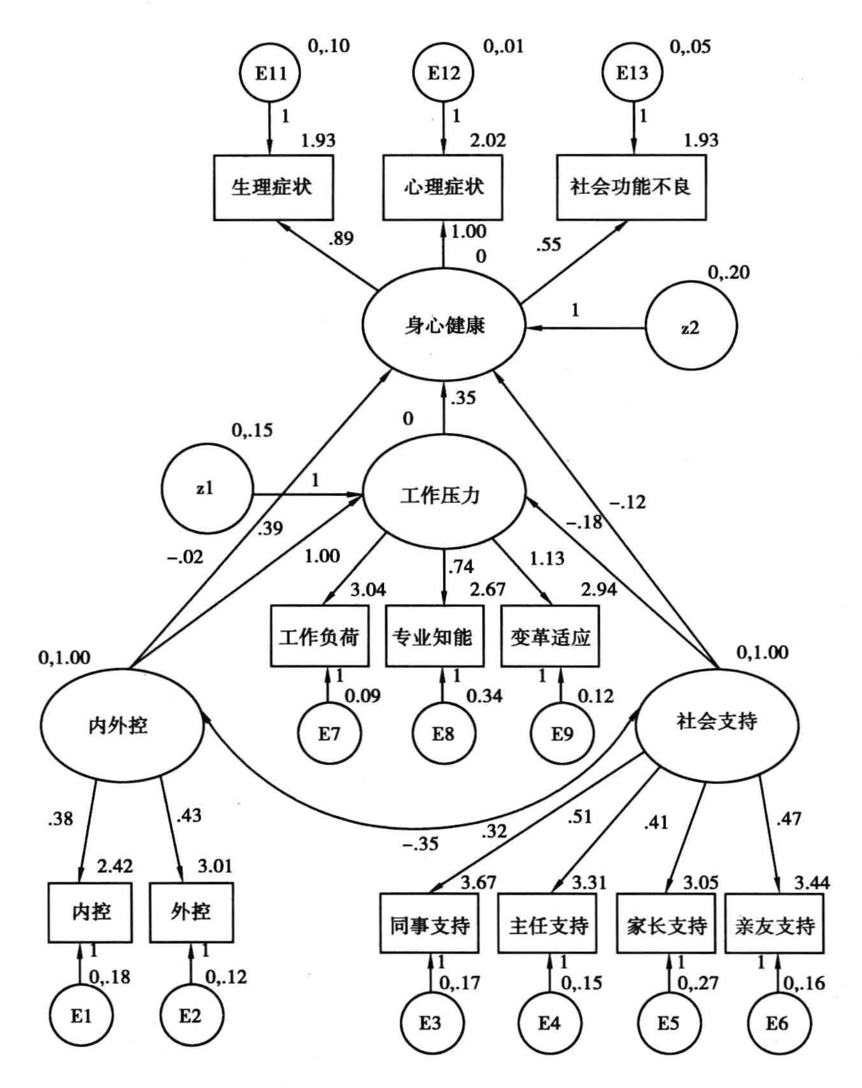


图 6-1　内外控信念、社会支持、教师工作压力与身心健康的关系模型

第一节　贝氏 SEM 参数估计误差与抽样样本大小的关系

吴丽华(2007)的原始抽取样本数为 630,在以下的模拟研究中将被视为研究的总体,以研究贝氏 SEM 参数估计误差与抽样样本大小的关系。笔者利用 SPSS 的数据的随机抽取程序,分别从原始总体(N = 630)中各抽取约 10%,20%,30% 的样本,作为 Bayesian SEM 分析的原始样本。SPSS 的数据随机抽取的程序,简单说明如下:

1. 点开 Data 下的[Select Cases];
2. 点选[Random Sample of Cases];
3. 点开[Sample];
4. 在[Approximately % of all cases]中输入抽取百分之多少,接着按下[继续]与[OK],即会在原始数据文件中产生一 个 filter_$(0/1)的变量;

5. 利用 SPSS 数据填补功能,填补缺失值(Amos bootstrapping 不允许数据含有缺失值);
6. 存文件。当您在 Amos 中利用 Data Files 调用此 SPSS 文件时,Amos 就会利用此 filter_$ 的变量,仅过滤出此 filter_$ 变量值为 1 的个案,纳入数据分析。

在进行 Bayesian SEM 时,必须先在分析属性窗口中 Estimation 下勾选[Estimate means & intercepts],接着打开[Analyze]表单下的[Bayesian Estimation]。点选它时,会出现如第二章图 2-1 的 Bayesian SEM 窗口,Amos 即会开始进行 MCMC 样本的产生。在执行 Bayesian SEM 时,如果 Bayesian SEM 经过很长一段时间仍然无法达到理想的聚敛标准(尤其是样本较小时),研究者可以按下暂停钮,试着修正估计参数的 Prior,尤其 CS 值较大者(如将 prior 由等概率分布改为正态分布),这样可加速各参数估计值的聚敛。以下将依样本的抽取比率逐一说明贝氏 SEM 参数估计误差与样本大小的关系。

一、抽出约 10% 的样本(n = 59)

在执行 Bayesian SEM 中,Amos 会自动提供 CS 统计量,以供诊断各参数的估计值是否已聚敛。CS 等于 1 表示完美聚敛(perfect convergence),此时更多的样本亦无法再提高估计统计量的精确性。因此,当 CS 逼近于 1.0 时,研究者就可终止样本的产生工作。通常如 CS 大于 1.1 时,则仍须继续产生更多的 MCMC 样本,以降低估计数的不确定性,而 Amos 则采取较保守的内定阈值 1.002。聚敛统计量(CS)小于 1.002 时,Bayesian SEM 窗口会出现笑脸,否则会出现红色的哭脸,在本例中 CS 值等于 1.0005,远小于内定阀值的 1.002。因此,在图 6-2 中窗口会出现笑脸。

图 6-2 左下方中(500 + 78369) * 128 表示 Amos 曾产生了 78869 * 128 个参数样本点,但只保留了 78369 个样本(式中的 * 号后的 128 表示 Amos 已进行了 7 次瘦身的工作($2^7 = 128$),因而仅保留了 1/128),式中 500 是内定的 burn-in 样本数(抽样分布未聚敛前所抛弃的样本数)。因为初期的样本向量通常不是取样于参数事后概率分布,Amos 会抛弃前 500 个样本,仅利用剩下来的 78369 个样本进行参数的事后概率的推估。

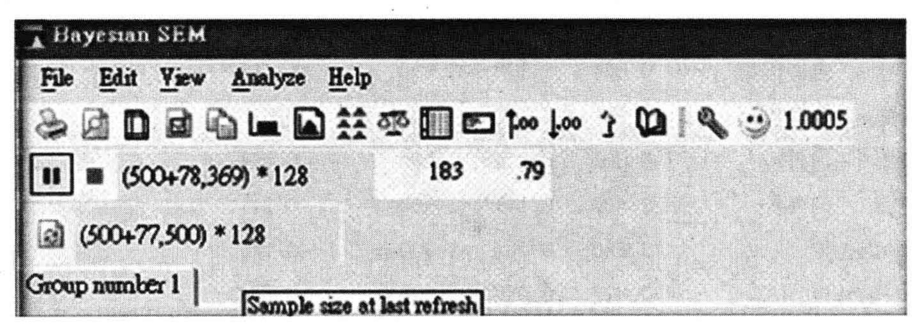


图 6-2 Bayesian SEM 分析窗口及 MCMC 相关的统计量

当您对 Bayesian SEM 分析结果满意时,即可按下图 6-2 的暂停钮,查看或打印如表 6-1 的各参数估计值摘要表。摘要表的内容输出,研究者可以自订。自订时,请开启 Bayesian SEM 分析表单中的 Options,即会出现图 6-3 的 Options 窗口,并在该窗口中点选所需的统计量数。表 6-1 的输出内容:Mean,Standard deviation,Convergence Statistics,0.95 Credible interval,即是勾选图 6-3 中选项的各项统计摘要表。

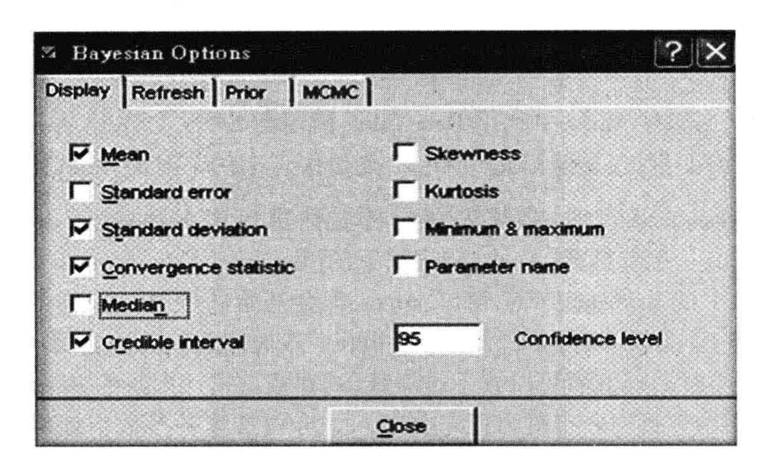


图 6-3　Bayesian Options 的设定窗口

表 6-1 中的 Mean，是当 N = 59 时由 78369 个 MCMC 样本点所计算出来的各参数估计值的平均数，S. D. 是各参数估计值的标准差，C. S. 则是各参数估计值的聚敛统计值，而 95% Lower bound 与 95% Upper bound 是各参数估计值的 0.95 上下置信区间。

表 6-1　Bayesian SEM 分析的各参数估计值的摘要表(N = 59)

Regression weights	Mean	S. D.	C. S.	95% Lower bound	95% Upper bound
家长支持 <--社会支持	0.445	0.103	1.000	0.251	0.656
主任支持 <--社会支持	0.543	0.094	1.000	0.365	0.736
同事支持 <--社会支持	0.335	0.080	1.000	0.181	0.497
亲友支持 <--社会支持	0.500	0.090	1.000	0.329	0.686
身心健康 <--工作压力	0.294	0.324	1.001	−0.364	0.894
专业知能 <--工作压力	0.689	0.138	1.000	0.420	0.961
身心健康 <--社会支持	−0.134	0.106	1.000	−0.345	0.069
社会功能不良 <--身心健康	0.547	0.067	1.000	0.418	0.680
生理症状 <--身心健康	0.881	0.100	1.000	0.688	1.083
工作压力 <--社会支持	−0.174	0.142	1.000	−0.412	0.130
身心健康 <--内外控	0.005	0.222	1.000	−0.419	0.482
工作压力 <--内外控	0.460	0.148	1.000	0.215	0.770
内控 <--内外控	0.380	0.094	1.000	0.198	0.569
外控 <--内外控	0.427	0.097	1.000	0.242	0.624

二、抽出约 20% 的样本(n = 138)

以下是当 N = 138 时，贝氏 SEM 统计分析的结果，执行过程与方法如同前述。在本例中 CS 值等于 1.0013，远小于 Amos 内定阈值的 1.002。因此，在图 6-4 右上方中窗口会出现笑脸。

图 6-4 左下方中，(500 + 63501) * 8 表示 Amos 曾产生了 64001 * 8 个参数样本点，

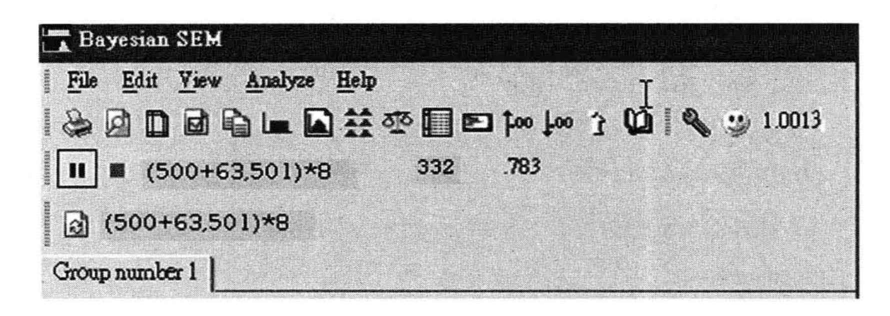


图 6-4　Bayesian SEM 分析窗口及 MCMC 相关的统计量

但只保留了 63501 个样本(式中的 * 8 表示 Amos 已进行了三次瘦身的工作,因而仅保留了 1/8),式中 500 是内定 burn-in 样本数(抽样分布未聚敛前所抛弃的样本数)。按下暂停钮后,研究者即可查看如表 6-2 的 Bayesian SEM 分析结果。

表 6-2　Bayesian SEM 分析的各参数估计值的摘要表(N = 138)

Regression weights	Mean	S. D.	C. S.	95% Lower bound	95% Upper bound
家长支持 <--社会支持	0.491	0.066	1.001	0.362	0.627
主任支持 <--社会支持	0.521	0.065	1.001	0.394	0.652
同事支持 <--社会支持	0.434	0.044	1.001	0.352	0.523
亲友支持 <--社会支持	0.466	0.048	1.000	0.374	0.562
身心健康 <--工作压力	0.300	0.175	1.001	-0.066	0.617
专业知能 <--工作压力	0.642	0.106	1.001	0.441	0.858
身心健康 <--社会支持	0.132	0.071	1.001	0.004	0.288
社会功能不良 <--身心健康	0.607	0.059	1.001	0.495	0.731
生理症状 <--身心健康	0.933	0.103	1.001	0.737	1.142
工作压力 <--社会支持	-0.077	0.078	1.001	-0.220	0.084
身心健康 <--内外控	0.312	0.111	1.001	0.122	0.560
工作压力 <--内外控	0.312	0.085	1.000	0.153	0.488
内控 <--内外控	0.394	0.061	1.001	0.280	0.519
外控 <--内外控	0.395	0.060	1.001	0.282	0.514

表 6-2 中的 Mean,是当 N = 138 时,由 63501 个 MCMC 样本点所计算出来的各参数估计值的平均数,其中 S. D. 是各参数估计值的标准差,C. S. 则是各参数估计值的聚敛统计值,而 95% Lower bound 与 95% Upper bound 是各参数估计值 0.95 的上下置信区间。

三、抽出约 30% 的样本(N = 201)

在本例中,N = 201,Bayesian SEM 分析窗口中的 CS 值等于 1.0002,远小于 Amos 内定阈值的 1.002。因此,在图 6-5 右上方中窗口会出现笑脸。

图 6-5 左下方中,(500 + 60071) * 256 表示 Amos 曾产生了 60571 * 256 个参数样本点,但只保留了 60071 个样本(式中的 * 256 表示 Amos 已进行了八次瘦身的工作,因而仅保留了 1/256),式中 500 是内定 burn-in 样本数(抽样分布未聚敛前所抛弃的样本数)。

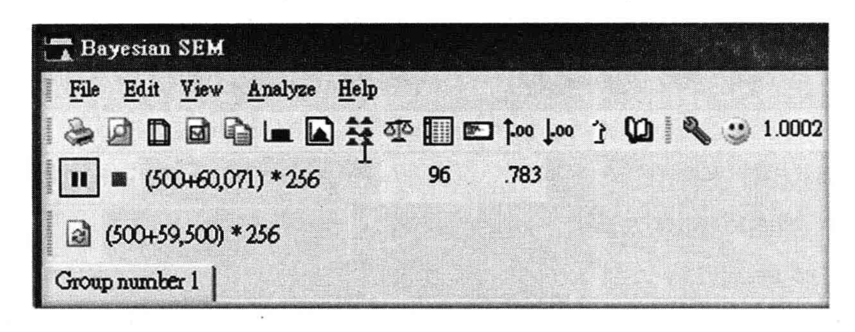


图 6-5 Bayesian SEM 分析窗口及 MCMC 相关的统计量

表 6-3 是当 N = 201 时，Bayesian SEM 的统计分析结果。

表 6-3 Bayesian SEM 分析的各参数估计值的摘要表（N = 201）

Regression weights	Mean	S. D.	C. S.	95% Lower bound	95% Upper bound
家长支持 <--社会支持	0.478	0.052	1.000	0.377	0.582
主任支持 <--社会支持	0.484	0.054	1.000	0.379	0.593
同事支持 <--社会支持	0.394	0.040	1.000	0.317	0.473
亲友支持 <--社会支持	0.399	0.043	1.000	0.316	0.485
身心健康 <--工作压力	0.230	0.195	1.000	-0.198	0.514
专业知能 <--工作压力	0.726	0.088	1.000	0.555	0.903
身心健康 <--社会支持	-0.004	0.062	1.000	-0.110	0.130
社会功能不良 <--身心健康	0.580	0.053	1.000	0.479	0.688
生理症状 <--身心健康	0.877	0.088	1.000	0.708	1.055
工作压力 <--社会支持	-0.017	0.068	1.000	-0.137	0.134
身心健康 <--内外控	0.228	0.129	1.000	0.044	0.520
工作压力 <--内外控	0.379	0.073	1.000	0.250	0.540
内控 <--内外控	0.346	0.049	1.000	0.250	0.444
外控 <--内外控	0.413	0.053	1.000	0.311	0.518

表 6-3 中的 Mean，是由当 N = 201 时，78369 个 MCMC 样本点所计算出来的各参数估计值的平均数，其中 S. D. 是各参数估计值的标准差，C. S. 则是各参数估计值的聚敛统计值，而 95% Lower bound 与 95% Upper bound 是各参数估计值的 0.95 上下置信区间。

图 6-6 中的数据是在三种样本大小（N = 59，N = 138，N = 201）下，各参数的估计平均值（B. D 栏位）、各参数的总体数（N = 630，E 栏），及各参数的估计误差（F. H 栏位），这些数据将供 SPSS 制作图 6-7 与图 6-8 时使用。

图 6-6 中变量 B-E，表示相关参数估计值在样本 N = 59 与总体 N = 630 上的差异值，其余变量 C-E 与 D-E，同此定义。由图 6-6 底部最后一行三个不同平均误差来看，整体的，三个参数平均差均在 0.081 以下，而 Bayesian SEM 参数估计误差的平均差随着样本的增大而逐渐变小。

估計參數	N=59	N=138	N=201	N=630	B-E	C-E	D-E
家長支持<--社會支持	0.445	0.49	0.478	0.4546	0.0096	0.0354	0.0234
主任支持<--社會支持	0.543	0.521	0.484	0.4825	0.0605	0.0385	0.0015
同事支持<--社會支持	0.335	0.435	0.394	0.4144	0.0794	0.0206	0.0204
親友支持<--社會支持	0.5	0.462	0.399	0.4603	0.0397	0.0017	0.0613
身心健康<--工作壓力	0.294	0.287	0.23	0.2512	0.0428	0.0358	0.0212
專業知能<--工作壓力	0.689	0.647	0.726	0.8396	0.1506	0.1926	0.1136
身心健康<--社會支持	-0.134	0.132	-0.004	-0.0209	0.1131	0.1529	0.0169
社會功能不良<--身心健康	0.547	0.603	0.58	0.5551	0.0081	0.0479	0.0249
生理症狀<--身心健康	0.881	0.928	0.877	0.9665	0.0855	0.0385	0.0895
工作壓力<--社會支持	-0.174	-0.074	-0.017	0.0046	0.1786	0.0786	0.0216
身心健康<--內外控	0.005	0.316	0.228	0.1854	0.1804	0.1306	0.0426
工作壓力<--內外控	0.46	0.32	0.379	0.3284	0.1316	0.0084	0.0506
內控<--內外控	0.38	0.394	0.346	0.3174	0.0626	0.0766	0.0286
外控<--內外控	0.427	0.397	0.413	0.3604	0.0666	0.0366	0.0526
內外控<-->社會支持	-0.377	-0.503	-0.462	-0.3825	0.0055	0.1205	0.0795
平均差					0.080973	0.046096	0.025043

图 6-6　贝氏 SEM 参数估计平均值及其在不同样本上的平均差

另外,由图 6-7 中三个不同样本的误差曲线整体来看,N = 201 的曲线所代表的各参数估计值的误差(与总体差距)最小,而 N = 59 的曲线所代表的各参数估计值的误差(与总体的差距)最大,亦显现出样本越大其参数的估计误差则越小。这意味着,虽然 Bayesian SEM 较适合于小样本的数据分析,但是所使用的样本仍然不可过小。值得注意的是,工作压力对专业知能的参数估计误差出现异常的形态,而且其参数估计误差亦最大。

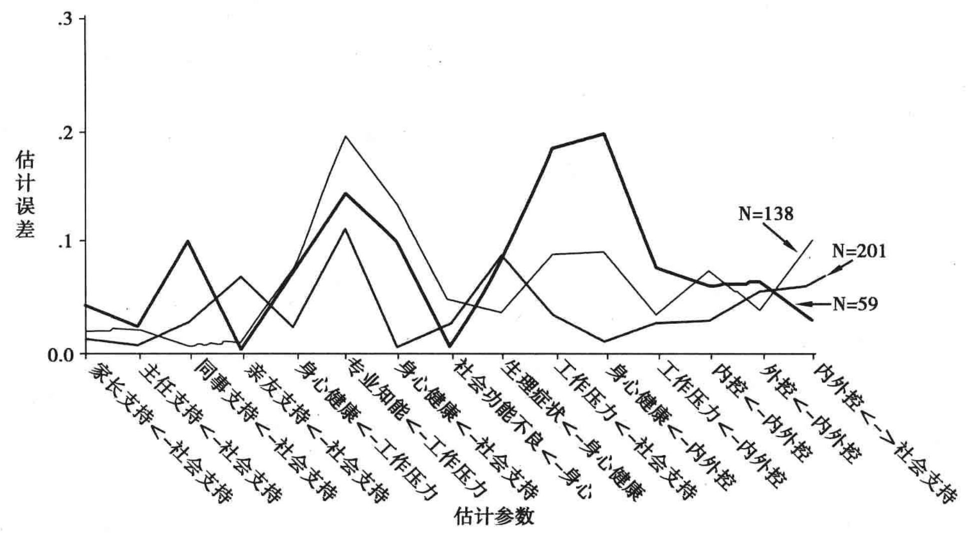


图 6-7　贝氏 SEM 参数的估计误差与样本大小关系的线形图

研究者如欲将图 6-7 的线形图转换成条图,亦可利用 SPSS 的图形编辑器或条图定义窗口,即可转换成图 6-8 的贝氏 SEM 参数估计误差与样本大小关系的条图。读者从此条图,当能更清楚地分辨出各参数在不同样本上的估计误差大小。一般来说,B-E(N = 59)的条件下,估计误差最大,而 D-E(N =201)的条件下,估计误差最小。

前述图 6-7 的线形图与图 6-8 的条图,是利用 SPSS 的统计图中的线形图制作出来

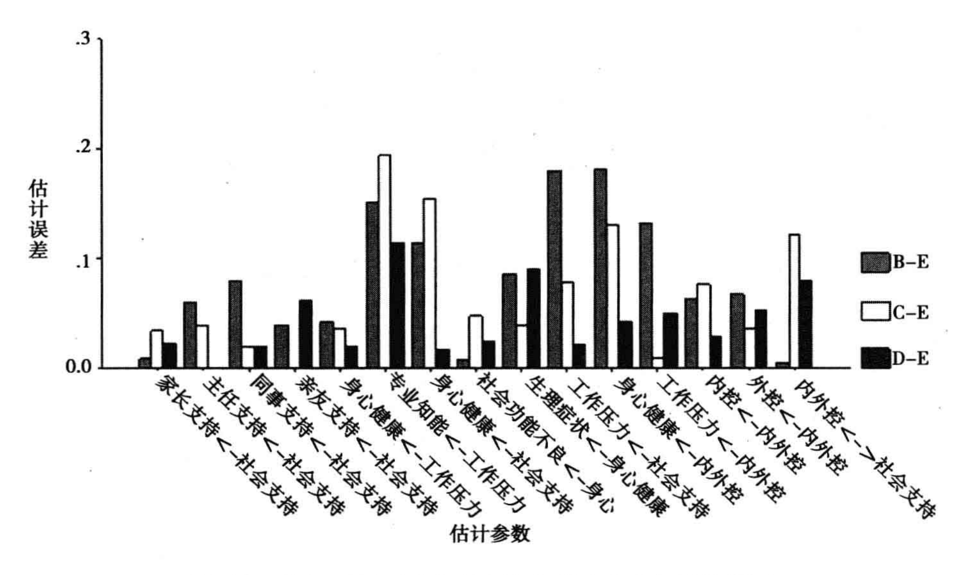


图 6-8 贝氏 SEM 参数估计误差与样本大小关系的条图

的,制作的步骤简述如下:首先,点选如图 6-9 的[复线图],并点选个别观察值数值。接着按下定义按钮,即会出现图 6-10 线形图的定义窗口。

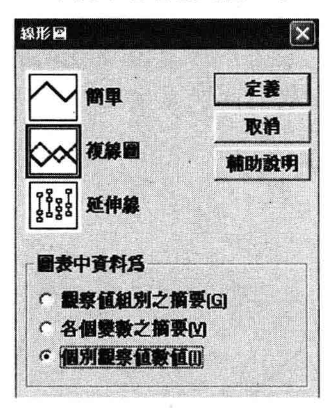


图 6-9 线形图形态的选择窗口

在图 6-10 线形图的定义窗口中,研究者须将曲线所欲代表的变量,输入线形图表示窗口中,并输入类别轴变量,最后按下确定。

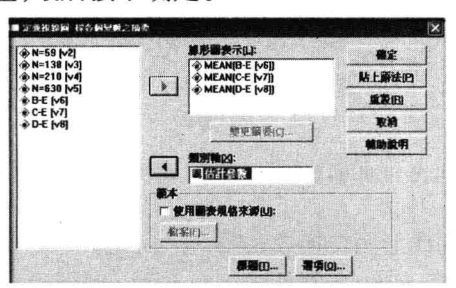


图 6-10 线形图的定义窗口

第二节 Bootstrapping 的参数估计误差与抽样样本大小的关系

以下的模拟研究,旨在研究 Bootstrapping 参数估计误差与抽样样本大小的关系。如同前述,仍利用 SPSS 的数据随机抽取程序,分别从原始的总体(N = 630)中各抽取 10%,20%,30% 的样本,作为 Bootstrapping 分析的原始样本。进行 Bootstrapping 时,必须先在图 6-11 的分析属性窗口中 Bootstrap 下勾选[Perform Bootstrap]、[Bias-corrected Confidence Intervals](并输入置信水平 95%)与[Bootstrap ML],并输入 Number of bootstrap samples,接着点选[Analyze]表单下的[Calculate Estimates]。

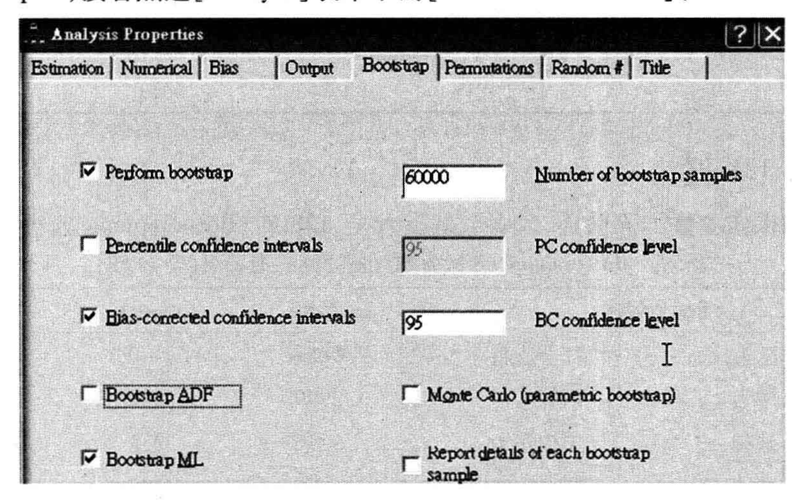


图 6-11 Amos Bootstrapping 的相关设定

一、当 N = 59 时

Amos Bootstrapping 的分析,在分析的属性窗口中设定完后,按照正常的 SEM 执行程序,即可获得如表 6-4 的分析结果。

表 6-4 Bootstrapping 分析的各参数估计值的摘要表(N = 59)

Parameter			Estimate	Lower	Upper
工作压力	<---	社会支持	-.1807	-.4068	.1071
工作压力	<---	内外控	.4047	.0524	.6779
身心健康	<---	工作压力	.3248	-1.0533	1.1371
身心健康	<---	社会支持	-.1218	-.4699	.1273
身心健康	<---	内外控	-.0116	-.6525	.6080
心理症状	<---	身心健康	1.0000	1.0000	1.0000
变革适应	<---	工作压力	1.0000	1.0000	1.0000
工作负荷	<---	工作压力	1.0000	1.0000	1.0000
家长支持	<---	社会支持	.4112	.1985	.7410

续表

Parameter			Estimate	Lower	Upper
主任支持	<---	社会支持	.5071	.2465	.6973
同事支持	<---	社会支持	.3135	.0576	.4839
亲友支持	<---	社会支持	.4635	.2238	.6630
专业知能	<---	工作压力	.6961	.2763	1.0996
社会功能不良	<---	身心健康	.5499	.3950	.7111
生理症状	<---	身心健康	.8842	.7151	1.1313
内控	<---	内外控	.3793	.1769	.5669
外控	<---	内外控	.4247	.2710	.5963
内外控	<-->	社会支持	-.3534	-.6650	.0874

二、当 N = 138 时

Amos 的操作方法如同前述，表 6-5 是当 N = 138 时，Bootstrapping 分析的结果。

表 6-5 Bootstrapping 分析各参数估计值的摘要表（N = 59）

Parameter			Estimate	Lower	Upper
工作压力	<---	社会支持	-.0857	-.2410	.0804
工作压力	<---	内外控	.2949	.1336	.4685
身心健康	<---	工作压力	.3290	-.1277	.6870
身心健康	<---	社会支持	.1134	-.0133	.2854
身心健康	<---	内外控	.2778	.1064	.6056
心理症状	<---	身心健康	1.0000	1.0000	1.0000
变革适应	<---	工作压力	1.0000	1.0000	1.0000
工作负荷	<---	工作压力	1.0000	1.0000	1.0000
家长支持	<---	社会支持	.4751	.3566	.6006
主任支持	<---	社会支持	.5040	.3679	.6281
同事支持	<---	社会支持	.4211	.3303	.5098
亲友支持	<---	社会支持	.4490	.3485	.5368
专业知能	<---	工作压力	.6440	.4139	.9284
社会功能不良	<---	身心健康	.6030	.4524	.7568
生理症状	<---	身心健康	.9306	.7208	1.1743
内控	<---	内外控	.3929	.2617	.5108
外控	<---	内外控	.3970	.2583	.5269
内外控	<-->	社会支持	-.4865	-.7175	-.2128

三、当 N = 201 时

Amos 的操作方法如前，表 6-6 是当 N = 201 时，Bootstrapping 分析的结果。

表 6-6 Bootstrapping 分析的各参数估计值的摘要表(N = 201)

Parameter			Estimate	Lower	Upper
工作压力	<---	社会支持	-.0317	-.1497	.1127
工作压力	<---	内外控	.3556	.2312	.5028
身心健康	<---	工作压力	.2746	-.1374	.5489
身心健康	<---	社会支持	-.0159	-.1347	.1138
身心健康	<---	内外控	.1951	.0268	.4615
心理症状	<---	身心健康	1.0000	1.0000	1.0000
变革适应	<---	工作压力	1.0000	1.0000	1.0000
工作负荷	<---	工作压力	1.0000	1.0000	1.0000
家长支持	<---	社会支持	.4680	.3686	.5640
主任支持	<---	社会支持	.4748	.3420	.5864
同事支持	<---	社会支持	.3857	.2926	.4676
亲友支持	<---	社会支持	.3908	.3013	.4708
专业知能	<---	工作压力	.7263	.5428	.9183
社会功能不良	<---	身心健康	.5802	.4515	.7066
生理症状	<---	身心健康	.8776	.7144	.0511
内控	<---	内外控	.3467	.2366	.4510
外控	<---	内外控	.4164	.3059	.5234
内外控	<-->	社会支持	-.4452	-.6493	-.2048

图 6-12 中的数据是在三种样本大小(N = 59,N = 138,N = 201)下各参数的估计平均值(B ~ D 栏位)、各参数的总体(N = 630,E 栏)及各参数在三种样本大小上的估计误差(F ~ H 栏位),这些数据将供 SPSS 制作图 6-13 的 Bootstrapping 的估计误差与样本大小关系线形图。

由图 6-12 底部最后一行三个不同平均误差来看,三个参数平均差均在0.0795 以下,而 Bootstrapping 参数估计误差的平均差随着样本的增大,其参数估计误差的平均差则逐渐变小,图 6-13 的线形图亦显示出趋势:样本越大,误差越小。此种形态与前述的 Bayesian SEM 非常相似。

诚如 Marcoulides 与 Schumacker(1996)所言,Bootstrapping(或 Bayesian SEM)并非小样本问题的万灵丹,我们不能期望利用 Bootstrapping(或 Bayesian SEM)就能满意地解决样本过小而不具代表性的问题,适度且具代表性的样本大小仍是 SEM 分析时的基本要求。

	A	B	C	D	E	F	G	H
1	估計參數	N=59	N=138	N=201	N=630	B-E	C-E	D-E
2	家長支持<--社會支持	0.4112	0.4751	0.468	0.4546	0.0434	0.0205	0.0134
3	主任支持<--社會支持	0.5071	0.504	0.4748	0.4825	0.0246	0.0215	0.0077
4	同事支持<--社會支持	0.3135	0.4211	0.3857	0.4144	0.1009	0.0067	0.0287
5	親友支持<--社會支持	0.4635	0.449	0.3908	0.4603	0.0032	0.0113	0.0695
6	身心健康<--工作壓力	0.3248	0.329	0.2746	0.2512	0.0736	0.0778	0.0234
7	專業知能<--工作壓力	0.6961	0.644	0.7263	0.8396	0.1435	0.1956	0.1133
8	身心健康<--社會支持	-0.1218	0.1134	-0.0159	-0.0209	0.1009	0.1343	0.005
9	社會功能不良<--身心健康	0.5499	0.603	0.5802	0.5551	0.0052	0.0479	0.0251
10	生理症狀<--身心健康	0.8842	0.9306	0.8776	0.9665	0.0823	0.0359	0.0889
11	工作壓力<--社會支持	-0.1807	-0.0857	-0.0317	0.0046	0.1853	0.0903	0.0363
12	身心健康<--內外控	-0.0116	0.2778	0.1951	0.1854	0.197	0.0924	0.0097
13	工作壓力<--內外控	0.4047	0.2949	0.3556	0.3284	0.0763	0.0335	0.0272
14	內控<--內外控	0.3793	0.3929	0.3467	0.3174	0.0619	0.0755	0.0293
15	外控<--內外控	0.4247	0.397	0.4164	0.3604	0.0643	0.0366	0.056
16	內外控<->社會支持	-0.3534	-0.4865	-0.4452	-0.3825	0.0291	0.104	0.0627
17	平均差					0.079433	0.041439	0.025556
18								

图 6-12 Bootstrapping 参数估计平均值及其在不同样本上的平均差

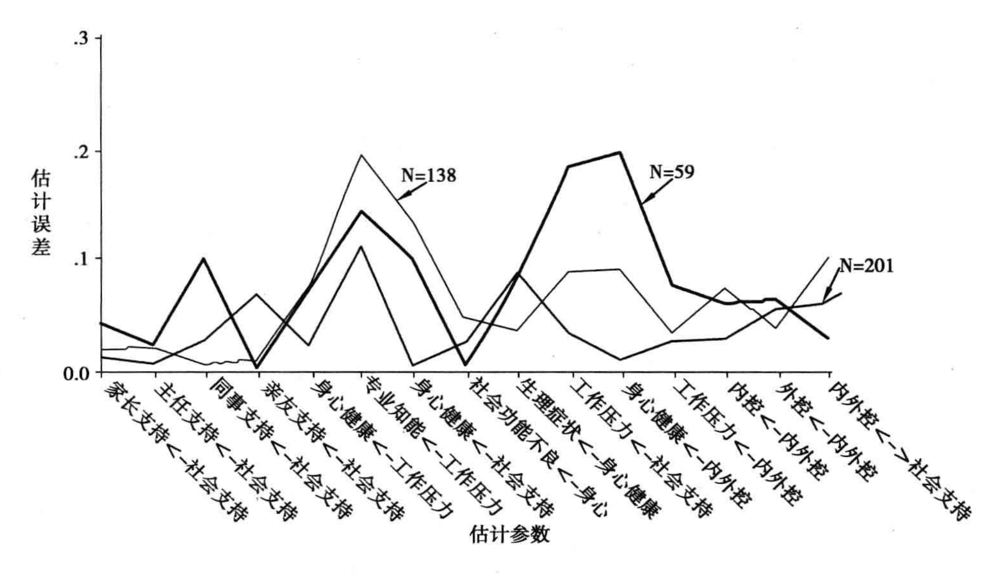


图 6-13 Bootstrapping 的估计误差与样本大小关系的线形图

第三节 传统 SEM 的参数估计误差与抽样样本大小的关系

图 6-14 中的数据,是传统 SEM 参数估计平均值及其在不同样本上的平均误差。

根据图 6-14 的数据,如图 6-15 所示,均反映出:在抽出样本为 20% 时,估计误差在工作压力对专业知能的估计参数上出现偏高现象;而在抽出样本为 10% 时,社会支持对工作压力的估计参数与内、外控对身心健康的估计参数上亦出现偏高现象,这可能反映出测量误差、受试者有无认真作答及抽样误差的问题。

	A	B	C	D	E	F	G	H
1	估計參數	N=59	N=138	N=201	N=630	B-E	C-E	D-E
2	家長支持<--社會支持	0.4147	0.4751	0.4691	0.4546	0.0399	0.0205	0.0145
3	主任支持<--社會支持	0.5115	0.504	0.476	0.4825	0.029	0.0215	0.0065
4	同事支持<--社會支持	0.3162	0.4211	0.3867	0.4144	0.0982	0.0067	0.0277
5	親友支持<--社會支持	0.4675	0.449	0.3918	0.4603	0.0072	0.0113	0.0685
6	身心健康<--工作壓力	0.3248	0.3296	0.2746	0.2512	0.0736	0.0784	0.0234
7	專業知能<--工作壓力	0.6961	0.644	0.7263	0.8396	0.1435	0.1956	0.1133
8	身心健康<--社會支持	-0.1228	0.114	-0.016	-0.0209	0.1019	0.1349	0.0049
9	社會功能不良<--身心健康	0.5499	0.6054	0.5802	0.5551	0.0052	0.0503	0.0251
10	生理症狀<--身心健康	0.8842	0.934	0.8776	0.9665	0.0823	0.0325	0.0889
11	工作壓力<--社會支持	-0.1923	-0.0857	-0.0317	0.0046	0.1969	0.0903	0.0363
12	身心健康<--內外控	-0.0117	0.2772	0.1956	0.1854	0.1971	0.0918	0.0102
13	工作壓力<--內外控	0.4082	0.2949	0.3565	0.3284	0.0798	0.0335	0.0281
14	內控<--內外控	0.3826	0.3929	0.3476	0.3174	0.0652	0.0755	0.0302
15	外控<--內外控	0.4283	0.3969	0.4174	0.3604	0.0679	0.0365	0.057
16	內外控<->社會支持	-0.3534	-0.4865	-0.4452	-0.3825	0.0291	0.104	0.0627
17	平均差					0.08112	0.04155	0.025507
18								

图 6-14　传统 SEM 参数估计平均值及其在不同样本上的平均差

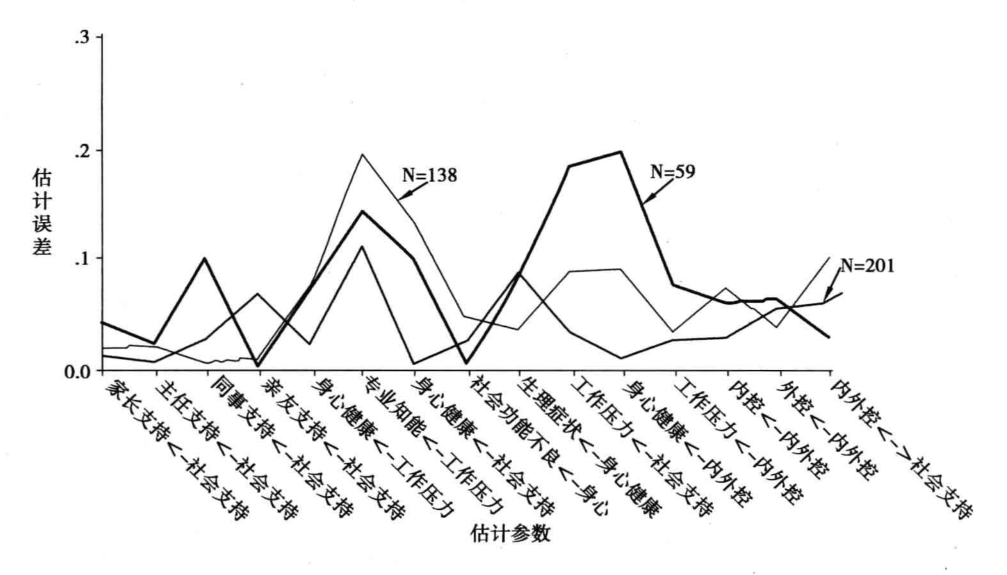


图 6-15　传统 SEM 参数估计误差与样本大小关系的线形图

第四节　贝氏 SEM 与 Bootstrapping SEM 的相对效能分析

表 6-7　三种统计方法在不同样本上的平均误差

	N = 59	N = 138	N = 210
传统 SEM	.08097	.04609	.02504
贝氏 SEM	.07943	.04144	.02556
Bootstrapping	.08122	.04155	.02551

根据图 6-7、图 6-13 与图 6-15 中的线形图比较分析知，三种统计方法的平均误差的形态相当类似，再由表 6-7 中的平均误差统计量来看，三种统计方法的平均误差的实质差异形态亦很类似。当样本逐渐增大时，三种统计方法的平均误差的分析结果趋于一致。因此，Bayesian SEM 与 Bootstrapping SEM 的效能不相上下，而且参数估计误差与抽样样本大小成反比，抽样百分比越大者，其参数的估计误差越小。这似乎意味着 Bayesian SEM 的应用虽然可以降低参数的估计误差，但仍然不能忽视样本大小的问题。换言之，由 Bayesian SEM 与传统 SEM 参数估计误差与样本大小关系的线形图来一一对比分析可以发现，不管样本大小，其各参数的曲线轮廓在 Bayesian SEM 与传统 SEM 上非常类似。原本在原始抽取样本上的参数估计误差较大者，在 Bayesian SEM 上，其参数估计误差仍然较大。例如，工作压力对专业知能的估计参数、社会支持对工作压力的估计参数与内外控对身心健康的估计参数。这反映出，测量工具的效能与受试者认真作答或原始抽样要具有代表性的重要性，利用 Bayesian SEM 或 Bootstrapping 的数据复制方法并无法有效改善抽样误差的缺陷。综合观之，Bayesian SEM 或 Bootstrapping 虽然被认为较适合于小样本的应用，以降低参数的计误差，但测量工具的信效度与样本大小、代表性仍然会严重影响参数的正确估计。尤其原始的抽取样本如果太小或不具有代表性，不管是 Bayesian SEM 或 Bootstrapping 似乎均无法有效改善参数估计值的正确性。

本章习题

1. 简述 Bayesian SEM 与 Bootstrapping SEM 的使用时机。
2. 比较说明 Bayesian SEM 与 Bootstrapping SEM 的相对效能。

第七章　复核效度的意义与应用实例

本章旨在介绍复核效度(cross-validation)及其在SEM上的应用分析。Hoyle与Panter(1995)曾呼吁在所有的SEM分析上,皆应进行效度复核的研究以检验模型的正确性、预测力与推论性,尤其是经过事后修正的理论模型。MacCallum等人(1992)针对心理学杂志上的SEM研究模型,经过修正后有无进行效度复核进行统计,结果发现鲜少研究者进行效度复核,而Holbert与Stephenson(2002)针对沟通科学的SEM论文评析,发现几乎没有研究者进行效度复核,这些研究发现值得其他领域的研究者警惕。

第一节　复核效度的意义与策略

复核效度是指测量的结果具有跨样本(或跨情境)的有效性,通常用来检验理论模型的稳定性、推论性及抽样的代表性。在结构方程模型中,一个理想模型在不同样本(通常来自同一总体)上出现的一致程度,即称之为模型的复核效度。假如该结构方程模型(尤其是测量模型)代表的是一个测量工具的因素结构,所得的结果即为测量工具的复核效度。由此观之,SEM复核效度类似于复制研究(replications)。复核效度的检验,依照MacCallum,Roznowski,Mar与Reith(1994)引自Bentler(1980)的定义,主要策略有三:

1. 宽松复核取向(loose replication strategy)

效度样本仅须维持与估计模型相同的因素结构设定,参数无须进行任何等同设定,所有的参数均开放估计。

2. 适中复核取向(moderate replication strategy)

指效度样本的部分参数必须套用估计样本的参数数据,也就是说效度模型中的部分参数必须进行样本间等同限制。此取向,一般惯用的策略亦有三种:

(1)固定因素负荷量,其余全部开放估计;

(2)固定加权系数,只开放估计方差或协方差;

(3)固定结构系数,只开放估计测量误差及残差之方差或协方差。

3. 严谨复核取向(tight replication strategy)

严谨复核策略不只两个样本之间具有完全相等的模型设定,同时对等参数的数值也完全设定为相等,无须进行任何参数的估计。严格来说,将测量误差及残差的方差或协方差进行等同的限制,常是不切实际的做法,因为各个样本间必然会有或多或少的变异。因此,固定因素负荷量与结构系数,是较务实的做法。

第二节 复核效度统计分析程序

根据 Cudeck 与 Browne(1983)的做法,SEM 复核效度的步骤主要有二:

首先,将样本切割为两个样本,利用多样本模型分析针对同一个假设模型进行估计,其中一个样本称为估计样本(calibration sample),另一个样本则为效度样本(validation sample)。假如这两个样本的角色互换,就是效度双复核(double cross-validation)的检验。其次,在估计样本上进行 SEM 模型参数的估计后,将所得的参数估计值,套入先前在估计样本上所使用的模型中(因此不须进行任何参数的估计),进行理论模型在效度样本上的检验。值得一提的是,此分析的适配差异函数值会等于复核效度指标(cross-validation index, CVI_k),其计算公式为:

$$CVI_k = F | S_b, \sum \hat{\lambda}_{k|a} |$$

$\sum \hat{\lambda}_{k|a}$ 表示 a 样本的参数估计值套入 k 模型的 b 样本中。因此,此 CVI_k 是一种双样本的 CVI 指标。

除了 CVI_k 指标的外,Browne 与 Cudeck(1989)又提出利用单一样本估计出来的 ECVI(Expected cross-validation index)指标,ECVI 指标是指所提议的模型在另一独立样本上的可能适配程度。ECVI 亦可用来评估 non-nested 模型,以选择较佳的模型,通常希望 ECVI 的值越小越好。Amos 亦提供此指标,其计算公式为 $ECVI = Fmin * (2q)/(N-1)$,q 代表待估计的参数数目,当 $q=0$ 时(不估计任何参数),ECVI = Fmin。由此公式可推知,影响复核效度的因素包含模型品质、因素负荷量大小、模型复杂度(参数个数)、样本大小与抽样代表性。例如,测量模型较佳时或简单的模型在较小样本上,其复核效度通常会较佳;而复杂的模型在较大的样本上,其复核效度会较佳(MacCallum, Roznowski, Mar & Reith, 1994; Whittaker & Stapleton, 2006)。通常复核效度指标的值会大于估计样本所得的差异函数值,因为估计样本的参数值解永远会比在效标样本上来得更适配。至于复核效度的统计方法与检验对象摘要如下,以供研究者参考:

表 7-1 复核效度的检验方法与对象

复核策略	检验对象	统计方法
宽松取向	因素结构	个别 SEM 分析、$\Delta\chi^2$ 分析
适中取向	结构 + 部分参数	多群组恒等性分析,ECVI 法
严谨取向	结构 + 全部参数	多群组恒等性分析,ECVI 法

第三节 复核效度的不同形式与内容

复核效度会因总体的异同与模型的多寡,而有不同的检验内涵(邱皓政,2004;黄芳铭,2004),参见表 7-2。

表 7-2　复核效度的不同形式与内容

模型数目	效度样本来源	
	相同总体	不同总体
单一模型	模型稳定性	效度延伸性
多重模型	模型选择性	效度类推性

前述 Cudeck 与 Browne(1983)的复核效度做法,即为模型稳定性(Model stability)的典型做法。这是一般研究者最感兴趣的,也是最基本的焦点。模型稳定性成立了,在不同总体上进行效度延伸性(Validity extension)检验才有意义。同样,多重模型选择性(Model selection)成立了,在不同总体上进行模型效度类推性(Validity generalization)检验才有意义。复核效度在多重模型的检验,旨在模型的选择,这是复核效度的另一用途。至于单一模型的统计分析方法,最常用的是 $\Delta\chi^2$ 分析,多重模型的统计分析方法,最常用的是多群组恒等性分析与 ECVI 比较法。

第四节　Amos 应用实例分析与解释

兹以吴丽华(2007)的硕士论文为例,说明如何进行测量工具的复核效度检验。检验的方法除了利用[Manage Models]窗口的参数设限外(具体操作步骤请参考笔者(2006)专著第九章),尚可利用以下两种较省力的方法,进行测量工具的复核效度检验。

第一种,首先建构理论模型如图 7-1 及利用[Manage Groups]窗口建立两个群组:估计样本与效度样本,如图 7-2。接着,移动鼠标指向任一待估计的路径(如教师工作压力与专业知能的路径对象),并按下鼠标右键,即会出现图 7-3 的对象属性表单,接着点选此对象表单,就会出现图 7-4 的对象属性窗口,在此窗口中的[Regression weight]下,标注路径名称如 p1,并点选图 7-4 的对象属性窗口中的[All groups],标注完毕后同法再继续标注其他参数路径的名称。点选对象属性窗口中的[All Groups],代表其他组的标签亦是仿此标签设定,因而可省去其他组参数标签等同的设定工作,请参见完成图 7-1。

第二种方法,是使用如图 7-5 的 Amos 多群组分析窗口,进行测量工具的复核效度检验,此种检验途径最为便捷,Amos 会依照您的模型设计,自动进行从最宽松到最严苛的组间参数等同的限制。

点选图 7-5 的多群组分析选目后,即会自动出现图 7-6 的多群组等同限制的设定窗口,可省去组间参数标签相同的设定,表中五个可能模型的内在设定,研究者亦可做增删。例如,研究者如不须进行测量误差等同的限制,就可将第 5 模型下的[✓]利用鼠标右键点去。按下 OK 后,Amos 即会开始进行多群组等同限制的分析。

为让读者先了解 TLH 模型在估计样本及效度样本上的适配情形,特先将第一种分析方法的 Unconstrained 模型所得结果,呈现于图 7-7 与图 7-8 路径中,并将相关的统计分析结果摘要如表 7-3。

e7 e8 e9 e11 e12 e13

1 1 1 1 1 1

工作负荷 专业知能 变革适应 生理症状 心理症状 社会功能不良

1 p1 p2 1 a1 a2

教师工作压力 d1 身心健康

1 d2 d3 1

z1 1 z2

内外控信念

c1 c2

内控 外控

1 1

e1 e2

图 7-1　TLH 模型的路径图与标签设定

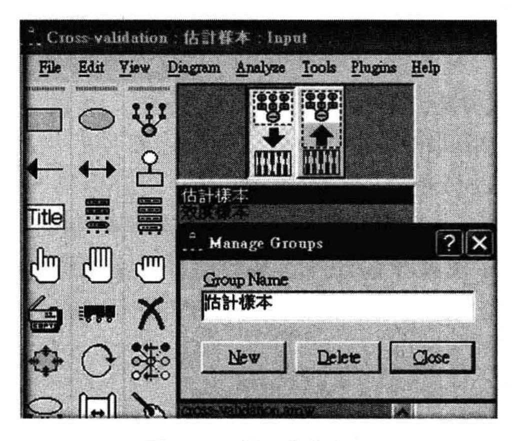


图 7-2　群组设定窗口

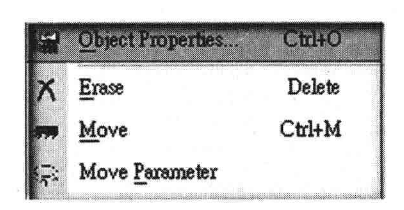


图 7-3 对象属性选目表单

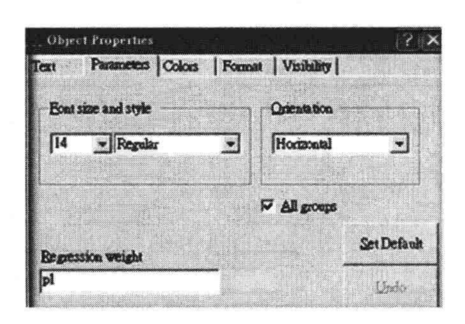


图 7-4 对象属性窗口

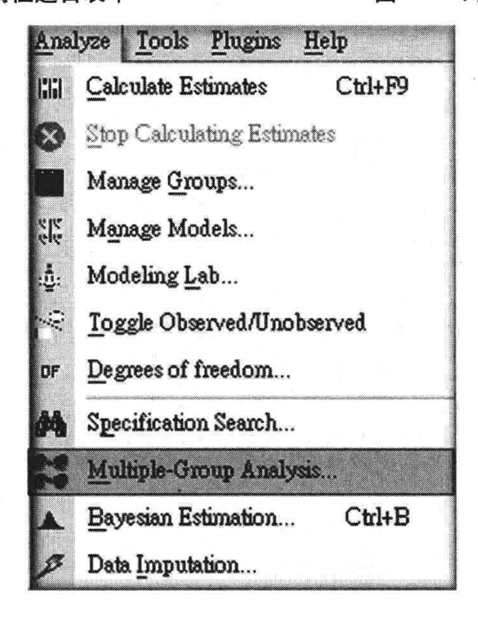


图 7-5 Amos 多群组分析选目

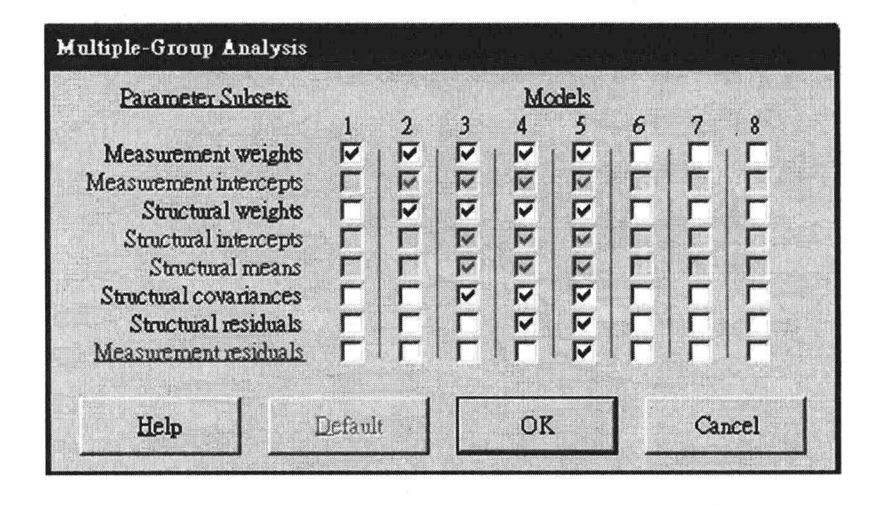


图 7-6 Amos 多群组分析的自动设定窗口

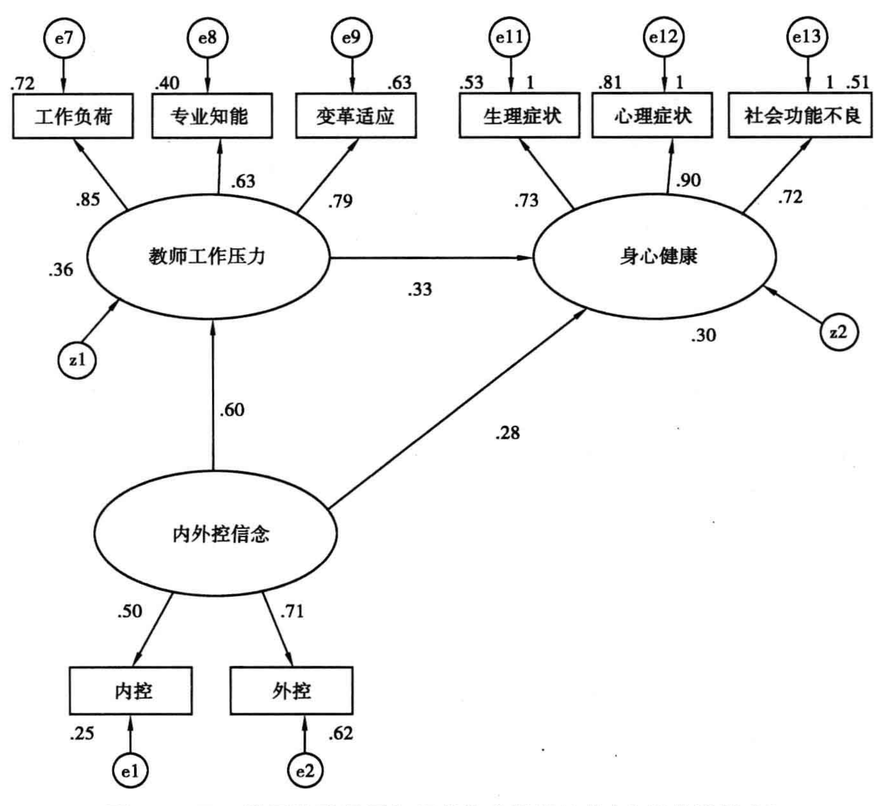


图 7-7 TLH 模型的路径图与标准化参数估计值(参数估计样本)

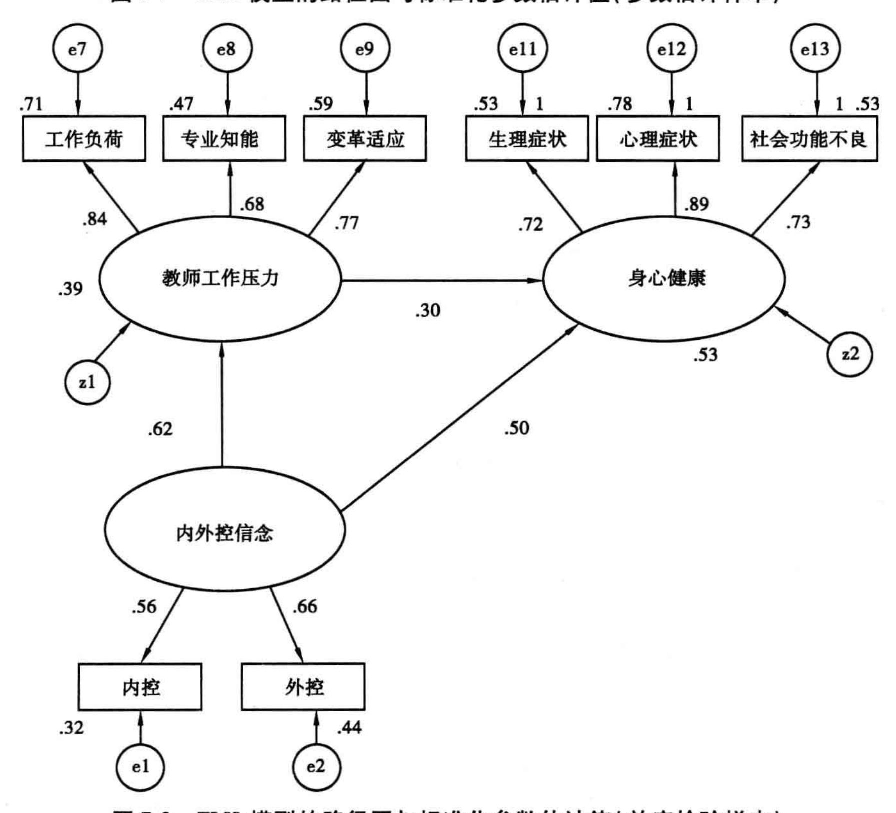


图 7-8 TLH 模型的路径图与标准化参数估计值(效度检验样本)

表 7-3　TLH 模型在参数估计样本与效度检验样本的整体模型适配度评鉴(N = 315)

指标名称		适配标准	参数估计样本	适配判断	效度检验样本	适配判断
绝对适配指标	χ^2	χ^2越小越好	19.828	是	32.582	否
		p 值至少≥.05	df = 17 p = .28		df = 17p = .01	
	GFI	GFI >.9 时,适配度佳	.984	是	.975	是
	AGFI	AGFI >.9 时,适配度佳	.967	是	.946	是
	SRMR	SRMR <.8 时,适配度佳	.030	是	.033	是
	RMSEA	RMSEA <.05 良好适配	.023	是	.054	是
		RMSEA <.08 可接受门槛				
	ECVI	ECVI 落入彼此的置信区间内,	ECVI = .184	是	ECVI = .225	否
		模型适配度佳	LO90 = .175		LO90 = .185	
			HI90 = .233		HI90 = .289	
增值适配指标	NFI	NFI≥.9 时,适配度佳	.977	是	.965	
	TLI	TLI≥.9 时,适配度佳	.994	是	.972	是
	CFI	CFI >.95 时,适配度佳	.997	是	.983	是
	RFI	RFI >.9 可接受	.962	是	.942	是
		RFI≥.95 时,适配度佳				
精简适配指标	PNFI	PNFI >.5 时,适配度佳	.593	是	.586	是
	PGFI	PGFI >.5 时,适配度佳	.465	否	.460	否
	AIC	AIC 越小,适配度佳	57.828	/	70.582	/
	CMIN/DF	CMIN/DF <1 过度适配	1.166	是	1.917	是
		CMIN/DF >3 模型适配度不佳,最好为 1 ~ 3				

由图 7-7、图 7-8 与表 7-3 的结果可知,不管“参数估计样本”还是“效度检验样本”,其路径系数及 χ^2值,从视觉判断来看,其实质上的差异并不大,相关的适配指标的数值实质上的差异也不大。基本上这是构念形态上的等同检验,旨在检验因素个数或形态在组间是否相同。如欲利用较严谨的统计方法,可根据 Cudeck 与 Browne(1983)的宽松检验做法,将所得的相对应参数估计值套入先前在估计样本上所使用的模型中(亦即手动键入参数估计值,请参见图 7-9),以检验在效度样本上的适配度。执行结果获得 F_{min} = 0.105,df = 36,χ^2 = 32.892,p = 0.617,并证实了 F_{min} = ECVI = 0.105。请注意图 7-9 的理论模型上,并没有任何参数需要重新估计,研究者旨在检验该理论模型的参数值在新样本上的适配度。

另外,由于估计样本数与效度样本数相同,因此研究者亦可以进行这两个样本的 χ^2 差异统计检验,分析结果得:

$\Delta\chi^2 = 32.892 - 19.828 = 13.064$,$\Delta df = 36 - 17 = 19$,$p > 0.05$。因此,两者的 χ^2 差异未达 0.05 的显著水平,显示出此 TLH 模型就因素形态上来说,不管从视觉上还是统计检验来说均具有稳定的复核效度。至于其测量模型及结构模型更严苛的不变性检验,亦可

利用前述的多群组分析(参见图 7-6),进一步检验分析如下。

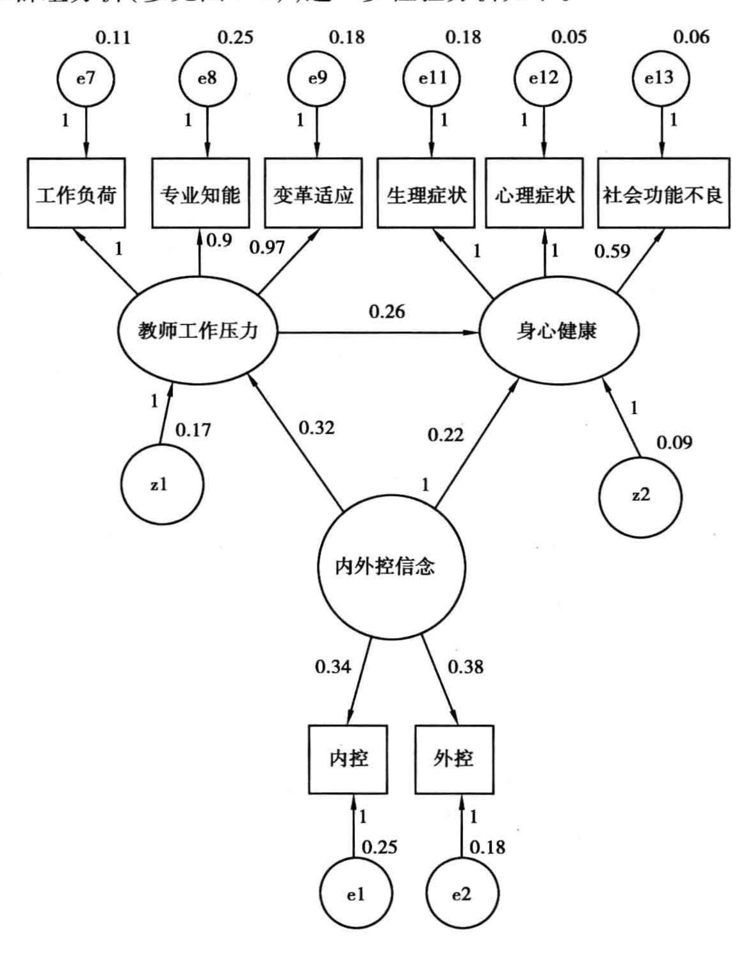


图 7-9 TLH 复核效度理论模型(未标准化估计值)

表 7-4 是为 Amos 多群组分析摘要表。由表 7-4 上半部各个模型的 p 值(均小于 0.05)看来,虽然并无法接受虚无假设(模型与数据相适配),但 RMSEA 值均小于 0.05, NFI 均大于 0.95,ECVI 值亦均落在彼此的 0.90 置信区间内。再从表 7-3 底部设限与未设限模型间的比较结果知,TLH 模型的参数估计样本与效度检验样本在构念形态(p = 0.424)、构念层次量尺(p = 0.398)、结构系数(p = 0.128)与残差方差上(p = 0.526),均具有组间不变性,而且 NFI 值亦均小于 0.05,符合 Little(1997)的建议标准。因此,TLH 模型的稳定性通过了从最宽松到最严苛条件的统计检验。

另外,由于 LISREL 在计算极小化函数及输入协方差矩阵是否为无偏估计值与 Amos 的处理方法略有不同,假如研究者欲使 Amos 与 LISREL 的分析结果相同,请进行以下的校正工作。

第一,在图 7-10 分析属性窗口[Estimation]内勾选[Emulisrel6],此项设定仅影响多群组的 SEM 分析,单群组分析完全不受影响,因为此时 χ^2 的计算为(N-G)F_{min}(G 表组别数)。由此观之,当您的极小化函数值 F_{min} 甚小时,[Emulisrel6]的影响力即不大了。

表 7-4 TLH 模型在参数估计样本(N=315)与效度检验样本(N=315)上的模型稳定性评鉴

模型	参数	卡方	自由度	p 值	RMSEA	NFI	ECVI (0.90 CI)
①未设限	38	52.41	34	.023	.029	.971	.204 (.179 ~ .242)
②因素负荷量相等	32	58.40	40	.030	.027	.968	.195 (.169 ~ .234)
③结构系数相等	29	61.36	43	.034	.026	.966	.190 (.163 ~ .230)
④结构残差相等	27	65.46	45	.025	.027	.964	.190 (.162 ~ .231)
⑤测量残差	19	72.56	53	.038	.024	.960	.176 (.147 ~ .218)
②-①最松		5.99	6	.424		.003	
③-②		2.96	3	.398		.002	
④-③		4.10	2	.128		.002	
⑤-④最苛		7.10	8	.526		.004	

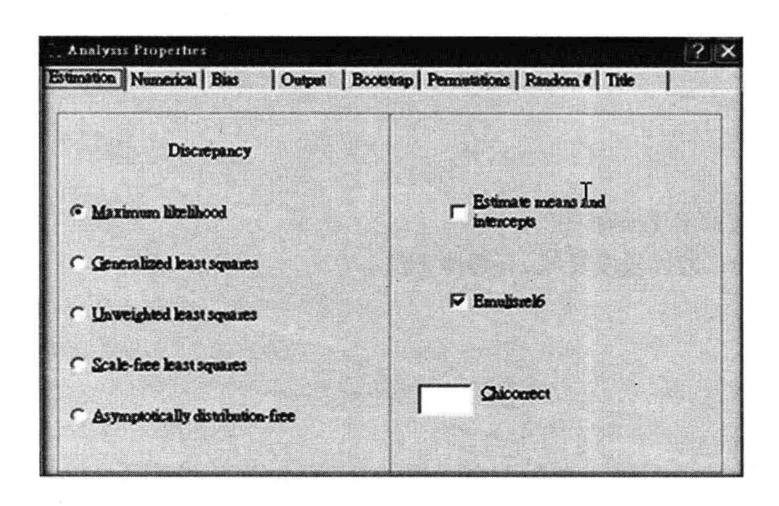


图 7-10 模拟 LISREL 的设定

第二,在图 7-11 分析属性窗口[Bias]下,Covariances supplied as input 与 Covariances to be analyzed 的小窗口内均勾选[Unbiased],亦即所输入的协方差矩阵与待分析的协方差矩阵均为无偏估计值(分母除以 N-1)。经过这两个步骤处理,Amos 与 LISREL 或 EQS 的分析结果将完全相同。不过,当样本逐渐增大时,即使不经上述校正,三个软件间的分析差异,亦将渐趋于和缓。

综合以上描述统计与推论统计来看,TLH 模型进行效度复核检验的结果,“参数估计样本”与“效度检验样本”在内在结构模型适配度复核所获得的参数估计值,和在整体模型适配度复核所获得的适配度评鉴指标,实质上的差异都不大,且再经 χ^2 差异值的统计

检验亦未达 0.05 的显著水平。因此，TLH 模型具有良好的复核效度，此模型相当稳定，具有可推论性。这亦反映出 TLH 的测量工具的表现极具稳定性，可有效应用于相同总体的其他样本，亦即反映了极佳的样本代表性。

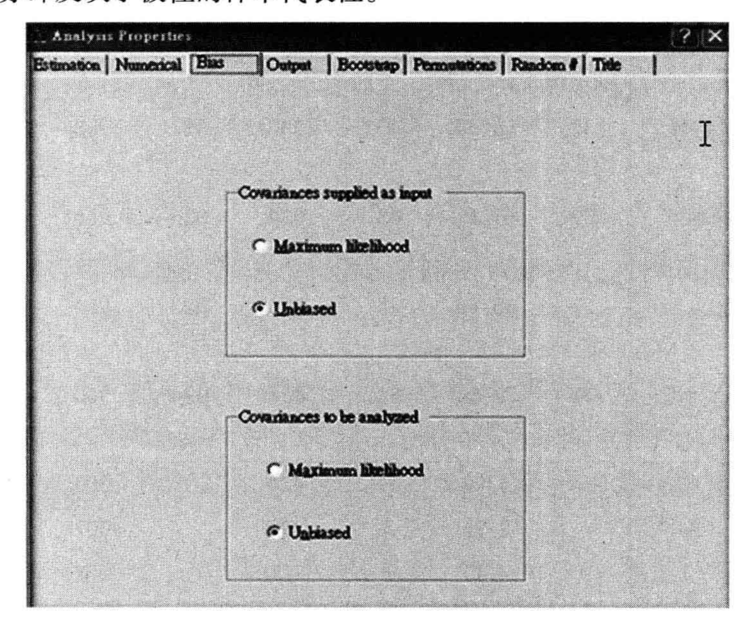


图 7-11　协方差矩阵的输入与分析属性的设定

本章习题

1. 效度复核的策略有哪些？
2. 利用 Amos 进行效度复核，有哪些方法？

第八章　回归分析与方差分析、SEM 分析的关系

因为方差分析是回归分析的特例，而回归分析又是 SEM 分析的特例，乃特辟此章分析回归分析与方差分析、SEM 分析间的关系，以利研究者利用 Amos 进行平均数结构分析或回归分析。不论利用回归分析进行传统方差分析(ANOVA)或 SEM 进行回归分析，如遇自变量或预测变量是类别变量，均须先将类别变量进行编码。以下将通过实例，说明与比较在回归分析或 SEM 分析中，类别变量编码(Coding)的三种主要方法，据此说明方差分析、回归分析、SEM 分析间的关系。

第一节　虚拟编码

虚拟编码(Dummy Coding)的方法是将所属类组编码为 1，非所属类组(作为控制组)编码为 0。最适用于有好几个实验组跟控制组的比较，各组样本大小相等或不相等均适用。以下先以林清山(2002，页 396)的三种教学方法(电视、编序、启发)的学习成绩数据为例，利用 SPSS 建构数据文件如图 8-1 所示：

	gourp	score
1	1.00	61.00
2	1.00	76.00
3	1.00	73.00
4	2.00	53.00
5	2.00	45.00
6	2.00	61.00
7	2.00	69.00
8	3.00	87.00
9	3.00	70.00
10	3.00	65.00

图 8-1　SPSS 所建构的数据文件

如果研究者利用传统的 ANOVA 分析的结果摘要如表 8-1，发现三种教学方法（电视、编序、启发）的学习成绩，并无显著差异（$F = 2.772, p = 0.130, \alpha = 0.05$）。

表 8-1　三种教学方法的方差分析摘要表

方差同质性检定

SCORE

Levene 统计量	分子自由度	分母自由度	显著性
.335	2	7	.726

方差分析

SCORE

	平方和	自由度	平均方和	F 检验	显著性
组间	564.000	2	282.000	2.772	.130
组内	712.000	7	101.714		
总和	1276.000	9			

表 8-2 则是利用虚拟编码格式，将图 8-1 的数据重新建构的新文件。针对 g 组类别变量来编码时，通常只需要 g-1 个虚拟变量，如表 8-2 中的 Group1 与 Group2。例如，最后一组启发组的编码，已可由 Group1 与 Group2 变量来定义。很显然，启发组的编码在 Group1 与 Group2 变量上均为 0，表示它是用来作为控制组。

表 8-2　虚拟编码格式的数据文件

GROUP	Y	Group1	Group2
电视组	61	1	0
	76	1	0
	73	1	0
编序组	53	0	1
	45	0	1
	61	0	1
	69	0	1
启发组	87	0	0
	70	0	0
	65	0	0

研究者如以表 8-2 的虚拟编码数据，以 Group1 与 Group2 变量为预测变量，以 Y 为效标，利用 SPSS 进行线性回归分析，可获得如表 8-3 的结果。分析的结果亦发现三种教学方法（电视、编序、启发）的学习成绩，并无显著差异（$F = 2.772, p = 0.130, \alpha = 0.05$）。读者不难发现方差分析与回归分析所得的结果与结论都完全相同。根据表 8-3 中 的未标准化系数，可建构以下的回归方程序：$Y = 74 - 4\text{Group1} - 17\text{Group2}$。式中，当 Group1 与 Group2 均为 0 时，$Y = 74$，是控制组：启发组的平均数。

根据表 8-2 的虚拟编码格式数据，就可建构出如图 8-2 的 Amos 回归路径图。接着，在进行文件连结时亦以此文件为连结标的。注意，进行 Amos 的回归分析时，必须先在分析属性窗口中勾选[Estimate means & intercepts]，以便进行平均数结构分析。

表 8-3　SPSS 回归分析的摘要表

方差分析[b]

模型		平方和	自由度	平均平方和	F 检验	显著性
1	回归	564.000	2	282.000	2.772	.130[a]
	残差	712.000	7	101.714		
	总和	1276.000	9			

a. 预测变量：(常数)，GROUP1，GROUP2

b. 依变量：SCORE

系数[a]

模型		未标准化系数		标准化系数		
		B 的估计值	标准误	Beta 分布	t	显著性
1	(常数)	74.000	5.823		12.709	.000
	GROUP2	−17.000	7.703	−.737	−2.207	.063
	GROUP1	−4.000	8235	−1.62	−.486	.642

a. 依变量：SCORE

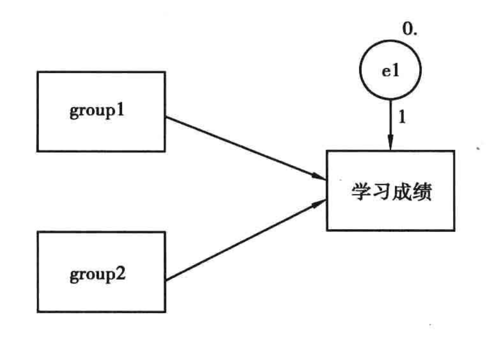


图 8-2　Amos 的三种教学方法回归路径设计图

Amos 的回归分析结果，如图 8-3 所示。利用此路径图中的相关参数估计值，就可建构出以下的回归方程：Y = 74 − 4Group1 − 17Group2。本例为三组间的比较，故只需两个虚拟变量，Group1 与 Group2。因第三组为控制组，上式中的截距“74”为控制组的平均数。Group1 的回归系数“ −4”为第一组平均数与控制组平均数的差(70 − 74)，此回归系数的 t 检验等于 SPSS 事后检验 LSD 法的检验结果；而 Group2 的回归系数“ −17”为第二组平均数与控制组平均数的差，因此，第二组的平均数为 57。又从图 8-4 中 Group1 与 Group2 的相关(−0.53)可推知，Group1 与 Group2 的多重比较是非正交比较。

另外，研究者可以利用图 8-4 中的 R^2(=0.442)进行 ANOVA 的 F 检验，计算公式及结果为：$F = \frac{0.442/df_1}{(1-0.44)/df_2} = \frac{0.442/2}{0.558/7} = 2.772$。再利用 SPSS 的内建函数，可以求出p =

1 – CDF. F(2.772, 2, 7) = 0.1298 ≈ 0.13。此项分析结果与前述 SPSS 的 ANOVA 分析结果完全相同。此项分析结果验证了方差分析是回归分析的特例。

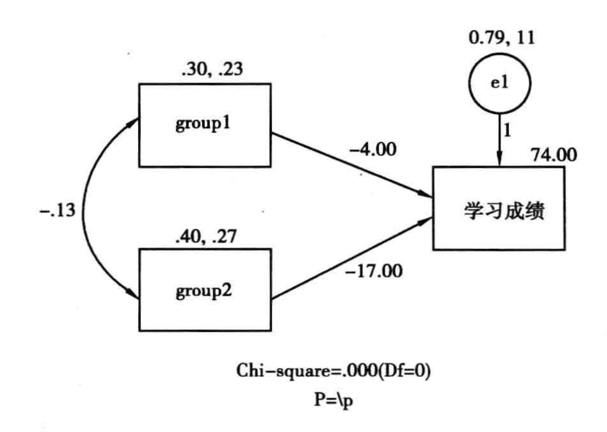


图 8-3　Amos 三种教学方法回归分析结果(未标准化)

请注意表 8-4 的回归系数的 CR 检验结果与表 8-3 底部的未标准化系数的 t 检验结果,两者略有出入(例如,第二组平均数与控制组平均数的差异检验结果在这两种方法上就完全不同:p = 0.063, t = –2.207; p = 0.0123, CR = –2.5025),主要原因是 Amos 使用最大概似法(ML)进行参数估计,而 SPSS 的回归分析是使用最小平方法(OLS)进行参数估计,因此导致不一致的标准误(SE)。不过,当样本趋近于无限大时,两者的分析结果应当趋于一致。

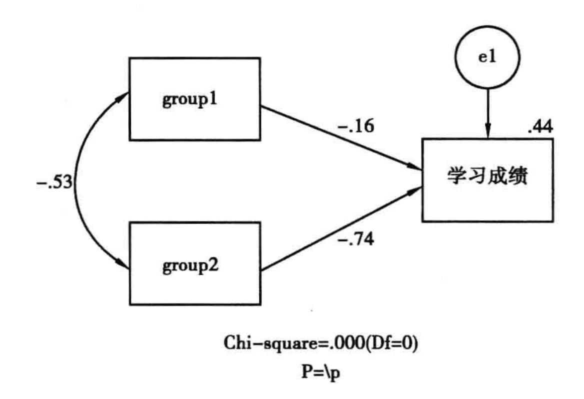


图 8-4　Amos 三种教学方法回归分析(标准化)

表 8-4　Amos 的回归系数检验

			Estimate	S. E.	C. R.	P
score	←	group1	–4.0000	7.2623	–.5508	.5818
score	←	group2	–17.0000	6.7932	–2.5025	.0123

以下为了检验 Group1 与 Group2 对于学习成绩的回归系数是否为 0(等同于在方差分析中,检验组间平均数是否相同的检验),乃将其相对应的回归系数设定为 0,参见图 8-5。检验结果亦发现三种教学方法(电视、编序、启发)的学习成绩,各平均数间并无显

著差异($\chi^2 = 5.251, df = 2, p = 0.072, \alpha = 0.05$)。将卡方值(5.251)除以 2 等于 2.63，这与方差分析摘要表中的 $F_{(2,7)} = 2.772$ 相差不大。当分子的自由度趋于无限大时，两者的数值及相对应的 p 值应合而为一(Li & Martin, 2002; Johnson & Balakrishnan, 1995)。由此可知，Amos 的平均数结构分析与传统的方差分析，当样本很大时，其统计分析结果并无二致。因为 χ^2 分布是 F 分布的极限分布，当分子的自由度趋近于无限大时(亦即样本很大时)，两者的分布趋近于一致。这两个分布的关系犹如正态分布与 t 分布的关系。这两个分布的关系可由右式推得：

$$\frac{\chi^2}{df} = F_{(df,\infty)} \text{ (Johnson \& Balakrishnan, 1995)}$$

式中 df 为分母的自由度，∞ 处为分子的自由度；更精确的两者对价关系请参看 Li 与 Martin(2002)文中的修正公式。由以上说明，可知方差分析、回归分析与 SEM 间具有不可分的关系，方差分析是回归分析的特例，而回归分析又是 SEM 分析的特例。不过，一般的研究常是使用有限的样本，χ^2 与 F 统计量间势必有些出入。事实上，在 SEM 上应用统计量时，样本亦须够大才能使 $(N-1)F_{ML}$ 或 $(N-1)F_{GLS}$ 逼近于 χ^2 分布。但 N 过大或过小(如 <50)亦会使 χ^2 的统计检验力过强或过弱(参见 p. 268; Bollen, 1989)。您可能已发现本研究(N = 10)的 χ^2 的 p 值(= 0.072)比 F 统计的 p 值(= 0.130)还小，即反映了此现象。由此观的，研究者在进行 SEM 分析时，进行适当的样本规划不可等闲视的，可参考应用本书所附的 SEM 样本规划的软件，明智地决定一个不会过大或过小的样本人数。

另外，请注意图 8-5 上学习成绩的截距 66$\left(= \frac{(70\times3)+(57\times4)+(74\times3)}{10}\right)$为加权总平均数。

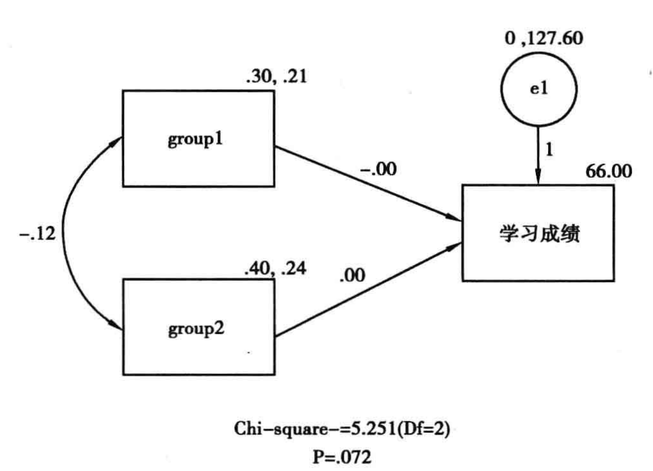


图 8-5　Amos 回归系数设定为 0 的检验

第二节 效果值编码

效果值编码(Effect Coding)是所属类组编为1,非所属类组编为0,但最后一组的编码为-1。效果值编码的使用最适用于无法决定何组为参照组,且欲进行正交或非正交多重比较,以反映出处理效果的大小。当运用于不等组时,方程序的截距等于各组平均数的未加权总平均数。兹将效果值编码格式列举于表8-5中。

表8-5 效果值编码格式的数据文件

Group	Y	Group1	Group2
电视组	61	1	0
	76	1	0
	73	1	0
编序组	53	0	1
	45	0	1
	61	0	1
	69	0	1
启发组	87	-1	-1
	70	-1	-1
	65	-1	-1

表8-6 SPSS回归分析摘要表

方差分析[b]

模型		平方和	自由度	平均平方和	F检验	显著性
1	回归	564.000	2	282.000	2.772	.130[a]
	残差	712.000	7	101.714		
	总和	1276.000	9			

a. 预测变量:(常数),GROUP2, GROUP1

b. 依变量:SCOFE

系数[a]

		未标准化系数		标准化系数		
模型		B的估计值	标准误	Beta分布	t	显著性
1	(常数)	67.00	3.219		20.816	.000
	GROUP1	3.000	4.654	.206	.645	.540
	GROUP2	-10.000	4.340	-.735	-2.304	.055

a. 依变量:SCORE

由表8-6底部的未标准化系数,可知 $Y = 67 + 3Group1 - 10Group2$。式中截距:“67”

$\left(=\dfrac{(70+57+74)}{10}\right)$为未加权总平均数。Group1 的回归系数"+3"为第一组平均数与总平均数的差(70-67),亦即第一组的处理效果;Group2 的回归系数"-10"为第二组平均数与总平均数的差(57-67),亦即第二组的处理效果。

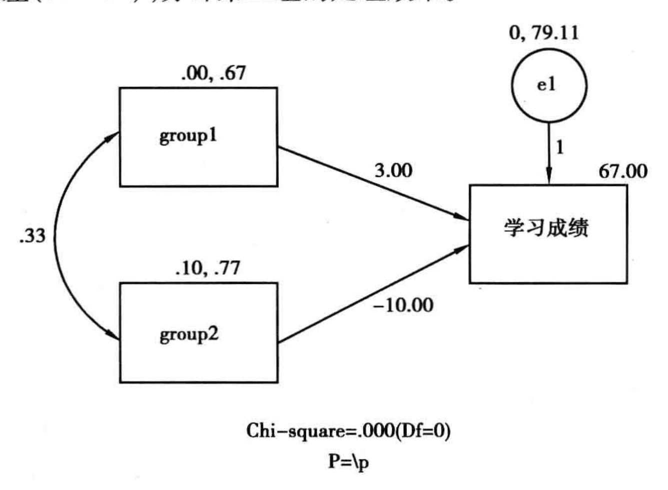


图 8-6　Amos 三种教学方法回归分析结果

本编码法,其回归系数反映出处理效果的大小,与前法的回归系数的意义不同。本例为三组间的比较,故只需两个效果值变量,Group1 与 Group2。又从图 8-6 中 Group1 与 Group2 的协方差(0.33)可推知,Group1 与 Group2 的多重比较是非正交比较。

表 8-7　Amos 的回归系数检验

			Estimate	S. E.	C. R.	P
score	←	group1	3.0000	4.1046	.7309	.4648
score	←	group2	-10.0000	3.8276	-2.6126	.0090

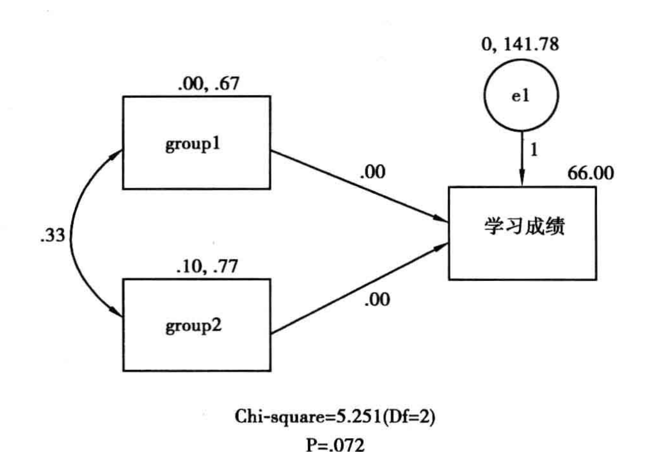


图 8-7　Amos 回归系数设定为 0 的检验

为了检验 Group1 与 Group2 对于学习成绩的回归系数是否为 0，乃将其相对应的回归系数设定为 0，参见图 8-7。检验结果亦发现三种教学方法（电视、编序、启发）的学习成绩，各平均数间并无显著差异（$\chi^2 = 5.251$，df = 2，p = 0.072，α = 0.05）。与前述图 8-5 中的检验结果相比，可以发现不管是使用虚拟编码或效果值编码，SEM 分析所得的 χ^2 值（= 5.251）与 p（= 0.072）值均完全相同。

第三节　正交编码

以表 8-8 中的 3 个组别为例，正交编码（Orthogonal Coding）是各对应元素的交乘积和为 0：$D = n_1C_1 + n_2C_2 + n_3C_3$。当不等组时，正交比较须符合 $\sum_{i=1}^{n} n_i C_{1i} C_{2i} = 0$，式中 n_i = 第 i 组人数，C_{1i} = 第一个比较中第 i 组的比较系数，C_{2i} = 第二个比较中第 i 组的比较系数。正交编码的适用时机：当研究者只对平均数的正交比较感兴趣时，而其回归系数的统计检验亦等于正交比较的显著性检验。因此，F 检验与组间平均数均等的检验可以同时进行。例如，研究者如对第一组平均数与第二组平均数的差异及对第一、二组加权平均数与第三组平均数间的差异感兴趣，可以考虑规划正交编码，此种格式的数据文件如表 8-8 所示。具体言之，第一对比较 D1 = (3)(4) + (4)(−3) + (3)(0) = 0，第二对比较 D2 = (3)(−12) + (4)(−12) + (3)(28) = 0；而这两对比较各对应元素的交乘积和亦等于 0：(3)(4)(−12) + (4)(−3)(−12) + (3)(0)(28) = 0，因此这两对比较为正交比较。

表 8-8　正交编码格式的数据文件

Group	Y	Group1	Group2
电视组	61	4	−12
	76	4	−12
	73	4	−12
编序组	53	−3	−12
	45	−3	−12
	61	−3	−12
	69	−3	−12
启发组	87	0	28
	70	0	28
	65	0	28

图 8-8 是利用正交编码的数据，进行 Amos 分析的结果。

由图 8-8 与表 8-9 底部的未标准化系数，可知 Y = 66 + 1.857Group1 + 0.286Group2。本例为三组间的比较，故只需两个编码变量，Group1 与 Group2。第一组平均数 70，可由 66 + 4 * 1.857 + (−12) * 0.286 求得。上式中截距："66"为加权总平均数。Group1 的回归系数" + 1.857"为第一组平均数与第二组平均数的比较（4 * 1.857 − (−3) * 1.857 = 13，为两组平均数的差），其 t 检验结果即为两组

平均数的差异检验结果。Group2 的回归系数"+0.286"为第一、二组加权平均数与第三组平均数的比较（$-12*.286-(28)*0.286=-11.44$）：$=\frac{70\times3+57\times4}{7}-74=62.57-74$。又从图 8-8 中 Group1 与 Group2 的相关（=0），再度验证了 Group1 与 Group2 多重比较是正交比较。

表 8-9　利用 SPSS 回归分析的摘要表

方差分析[b]

模型		平方和	自由度	平均平方和	F 检验	显著性
1	回归	564.000	2	282.000	2.772	0.130[a]
	残差	712.000	7	101.714		
	总和	1276.000	9			

a. 预测变量：（常数），GROUP2，GROUP1

b. 依变量：SCOFE

系数[a]

		未标准化系数		标准化系数		
模型		B 的估计值	标准误	Beta 分布	t	显著性
1	（常数）	66.00	3.189		20.694	.000
	GROUP1	1.857	1.100	.476	1.688	.135
	GROUP2	.286	.174	.464	1.642	.145

a. 依变量：SCORE

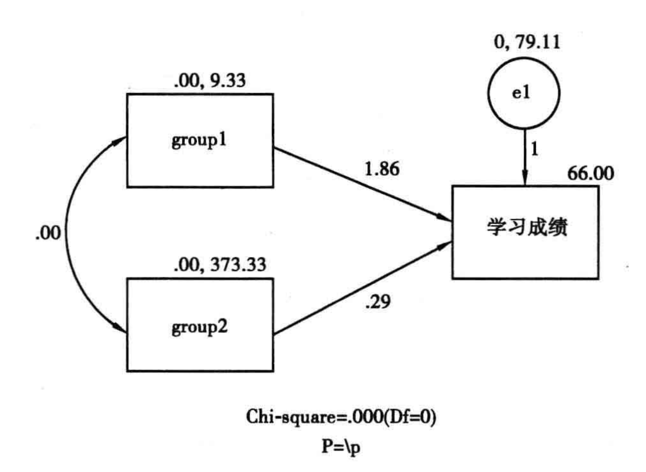


图 8-8　Amos 三种教学方法回归分析结果

表 8-10　Amos 的回归系数检验

			Estimate	S. E.	C. R.	P
score	←	group1	1.8571	.9705	1.9137	.0557
score	←	group2	.2857	.1534	1.8620	.0626

为了检验 Group1 与 Group2 对于学习成绩的回归系数是否为 0，乃将其相对应的回

归系数设定为0,参见图8-9。检验结果,是三种教学方法(电视、编序、启发)的学习成绩,各平均数间并无显著差异(χ^2 = 5.251,df = 2,p = 0.072,α = 0.05)。分析结果与前述图8-5 与图8-7 相比,可以发现不管是虚拟编码、效果值编码或正交编码方法,SEM 分析所得的χ^2值(= 5.251)与 p(= 0.072)值均完全相同。

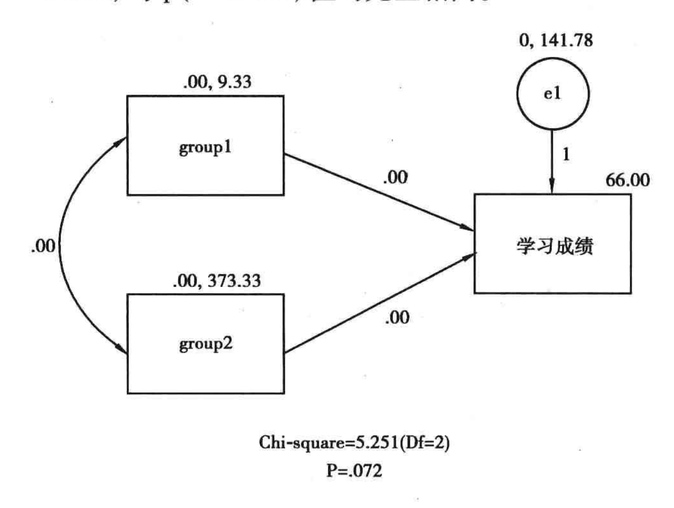


图8-9 Amos 回归系数设定为0 的检验

由以上的分析结果与关系可推知以下结论:

1. 不管是使用虚拟编码、效果值编码或正交编码方法,SEM 分析所得的χ^2值(= 5.251)与 p(= 0.072)值均相同。

2. χ^2与 F 分布间具有密切关系:χ^2分布是 F 分布的极限分布,当分子的自由度趋近于无限大时,两者的分布趋近于一致。这两个分布的关系犹如正态分布与 t 分布的关系 $\chi^2 = df * F_{(df, \infty)}$,式中 df 为分母的自由度,∞ 处为分子的自由度。以虚拟编码的实例来看,将卡方值(5.251)除以2 = 2.63 与方差分析摘要表中的 $F_{(2, 7)}$ = 2.772 甚为接近,当分子的自由度趋于无限大时,两者的数值及相对应的 p 值应合而为一。由此可知,Amos 的平均数结构分析与传统的方差分析,当样本很大时,其统计分析结果并无二致。

3. Amos 进行参数估计时,通常使用最大概似法,而 SPSS 的回归分析使用最小平方法进行参数估计,因此可能导致不一致的标准误,而产生不一致的统计分析结果或结论,尤其当样本不大时。

4. 类别变量的编码方法,在回归分析与 SEM 分析中均扮演着非常关键的角色,毋庸置疑,它是学习 SEM 的必备知能。

本章习题

1. 以图8-1 中的数据为例,利用虚拟编码、效果值编码与正交编码方法,进行第二组与第三组间平均数的比较,他们的结论会完全相同吗?

2. 试着利用 Amos 进行传统的协方差分析(ANOVA),结果会有不同吗?

第九章　交互作用效果与间接效果的检验

研究者在探讨双变量间的关系时,如果忽视变量间的关系是否因另一重要变量而产生调节或中介作用,常会导致肤浅或错误的结论。因此,熟悉调节或中介作用的探究方法,乃是学术研究者的必备能力。欲进行交互作用效果与间接效果检验,须先厘清调节变量与中介变量的异同,这两个名词常被研究者混为一谈,以致产生错误的看法。

调节变量(Moderation variables,如图 9-1-(a)中的 X2)是一种起始变量,也是一种次级自变量,他会使得原本的自变量(X1)对于依变量(Y)的效果,因调节变量(X2)而产生交互作用现象(interact with the effect of an IV on the outcome DV),此种交互作用分析,通常使用回归分析检验交互作用项(如 X1 * X2 两个调节变量的乘积)。

中介变量(Mediating variables, X2)在图 9-1-(b)中则充当 X1 的果及 Y 的因,具有双重角色(act as the casual link or path between IV and DV)。由图右侧 9-1-(b)知,X1 对 Y 的关系具有直接效果(路径 c)与间接效果(经由 a,b 路径),将直接效果与间接效果加总在一起即为 X1 对 Y 的总效果。此种中介效果分析,通常使用净相关检验,利用阶层式回归分析,将中介变量先放入回归方程中检验,或使用 SEM 检验间接效果。

简言之,调节变量的分析,即为交互作用效果的分析,可以告诉我们“什么时候”自变量会影响依变量;而中介变量的分析,即为间接效果检验的分析,可以告诉我们“为什么”自变量会影响依变量或自变量“如何”影响依变量。

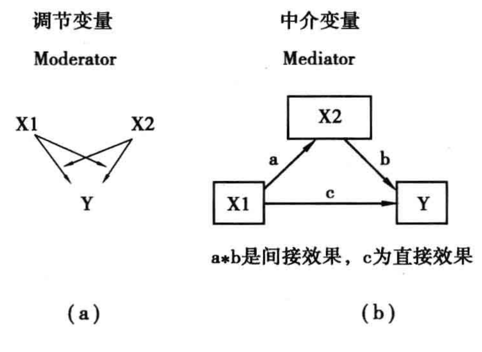


图 9-1　调节变量与中介变量的异同

以下本章节中,首先将讨论如何进行交互作用效果的检验,接着深入分析直接与间接效果的各种检验方法,最后论及条件式中介效果(conditional indirect effect)的检验方法,章末列举实例,印证说明各种中介效果的分析。

第一节　交互作用效果的检验方法

交互作用效果的统计检验，其方法较多，可以归纳如表 9-1（整理自 Jonsson，1998；Ping，1995 & 1996；Schumacker & Marcoulides，1998；陈顺宇，2007；余民宁，2006）。

表 9-1　交互作用效果的主要统计检验方法摘要表

变量属性	类　别	连　续
类　别	1. 传统的 two-way ANOVA 2. Dummy/Effect coding SEM 3. Dummy/Effect coding 回归法	1. 多群组 SEM 2. Dummy/Effect coding SEM 3. Dummy/Effect coding 回归法
连　续	1. 多群组 SEM 2. Dummy/Effect coding SEM 3. Dummy/Effect coding 回归法	1. 指标交乘积直接估计法（Indicant product approach）：Kenny & Judd 法 2. 两步骤法：Ping's 2-step 法与 Jonsson's 2-step 因素分数法

由表 9-1 可知，交互作用效果的统计检验方法，常须视变量的属性而定。调节变量或自变量如为类别变量，可先转换成虚拟/效果变量（Dummy/Effect variables），使用两个变量相乘积，以产生交互作用项。调节变量与自变量如皆为连续变量，须视交乘积项是观察变量还是潜在变量而定，如为潜在变量可采 Kenny 与 Judd（1984）的做法：增加另一潜在变量作为交互作用因子，该因子的测量指标为两个潜在变量下各观察指标的交乘积项。

另一可行方法是：先评估测量模型及计算两个潜在变量下各受试者的因素分数，接着计算各受试者在两个潜在变量下的因素分数乘积的交互作用因子的分数，最后进行这三个潜在变量的回归分析，这是 Jonsson's 双步骤（2-step）因素分量（Schumacker & Marcoulides，1998）。

至于 Ping's 双步骤（2-step）法，则可分为单指标法（single indicator approach）与多元指标法（multiple product indicator approach）。其双步骤的交互作用分析必须经过两个阶段的统计分析，第一阶段为测量模型的 CFA 分析，而第二阶段为全模型的 SEM 分析。遇到指标很多时多元指标法相当繁琐，首先须估计线性潜在变量指标的因素负荷量与误差变异量，利用线性潜在变量指标间的交乘积建立非线性指标，并利用第一阶段测量模型所得的估计值计算这些非线性指标的因素负荷量与误差变异量，以作为第二阶段结构模型中非线性指标的因素负荷量与误差变异量（作为常数）；最后，在第二阶段整个结构模型中，估计潜在变量间的线性与非线性关系（Ping，1996）。

当指标很多时，研究者可改用 Ping 的单指标法，本法在计算交互作用项时，先将各潜在变量下的指标加总成一新变量，接着利用两个加总变量的积作为交互作用项，最后进行模型分析与检验。此种模型分析，交互作用参数的估计途径也有两种：①交互作用项作为自由参数变量，直接进行参数估计，但本法目前只有 SAS Proc Calis 或 LISREL 8 以上才能执行；②交互作用作为固定参数变量，做法如同前述多元指标法，利用第一阶段测量

模型所得的估计值计算这些非线性指标的因素负荷量与误差变异量,以作为第二阶段结构模型中的固定常数(Ping, 1995)。

一般而言,当两个交互作用的变量均为类别变量时,使用传统的二因子(2-way)方差分析,应是最为便捷的。而在表9-1中其余的情境下,研究者会发现交乘积法进行交互作用分析不必分割样本,较为精简。但多群组SEM分析法较具弹性且可避免多元共线性问题及变量须具正态分布的基本假设,尤其当交互作用项是类别变量时,多群组SEM分析法更为适用。如果交互作用是潜在变量,因素分量应是最合理的选择(Joreskog, 1998)。

另外,为了避免多元共线性(multicollinearity)的发生与交互作用结构参数的偏估,笔者建议将所有自变量与依变量(交互作用项除外)化成离均差分数(mean-centering),或化成因素分数后,再进行交互作用分析。

因为交互作用效果的统计检验方法相当多,为节省篇幅,本书仅就虚拟变量、多群组SEM法、Jonsson's因素分量与Ping的单指标2-step法,各举一实例说明交互作用的分析方法。读者如欲了解更详细而具体的应用实例,请参阅Schumacker与Marcoulides(1998)的专书,该书中论及更多的交互作用分析方法。

一、虚拟/效果变量法

此种分析法较适合于分析两个变量中至少一个为类别变量,如果双变量均为连续变量,请参考Jonsson's因素分量与Ping的单指标2-step法,进行交互作用分析。

(一) 两个变量均为类别变量

本例将使用效果编码法(Effect coding,参见图9-2的高低压力组两个变量),进行二因子(gender与高低压力组2)方差分析,感兴趣的研究者亦可采用虚拟编码法进行分析。传统的方差分析结果如表9-2所示,显示出交互作用未达0.05的显著水平($F=0.327$, $p=0.568$)。

估計樣本315-出書.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

	gender	健康因素分數	高低壓力組	交互作用	高低壓力組2	交互作用2
1	1	-.05992	.00	.00	-1.00	-1.00
2	1	-1.02983	1.00	1.00	1.00	1.00
3	0	.01945	.00	.00	-1.00	.00
4	1	1.79838	1.00	1.00	1.00	1.00
5	0	-.26973	.00	.00	-1.00	.00
6	0	.25245	1.00	.00	1.00	.00
7	0	-.58446	.00	.00	-1.00	.00
8	1	-.84556	1.00	1.00	1.00	1.00
9	0	.54164	.00	.00	-1.00	.00
10	0	-1.02747	1.00	.00	1.00	.00
11	0	-.13909	1.00	.00	1.00	.00
12	0	-.47219	.00	.00	-1.00	.00

图9-2　SPSS5中gender & 高低压力变量交互作用分析的原始数据

由表9-2知,性别与高低压力的交互作用效果,并未达显著水平($p=0.568>\alpha=0.05$)。

研究者如改用Amos进行分析,此交互作用的路径图设计与分析结果,如图9-3与表9-3所示,显示出交互作用亦未达0.05的显著水平($CR=0.5744$, $p=0.5657$)。

表 9-2 传统二因子方差分析摘要表 Dependent Variable：健康因素分数

Source	Type Ⅲ Sum of Squares	df	Mean Square	F	Sig.
Correct Model	45.442ª	3	15.147	17.541	.000
Intercept	.117	1	.117	.135	.714
gender	.081	1	.081	.093	.760
高低压力组 2	36.585	1	36.585	42.367	.000
gender * 高低压力组 2	.282	1	.282	.327	.568
Error	268.558	311	.864		
Total	314.00	315			
Corrected Total	314.000	314			

a. R Squared = .145 (Adjusted S Squared = .136).

图 9-3 的分析模型是“Effect coding”SEM 的分析方法，本质上它是线性回归分析，相关的路径系数检验，请参见表 9-3。

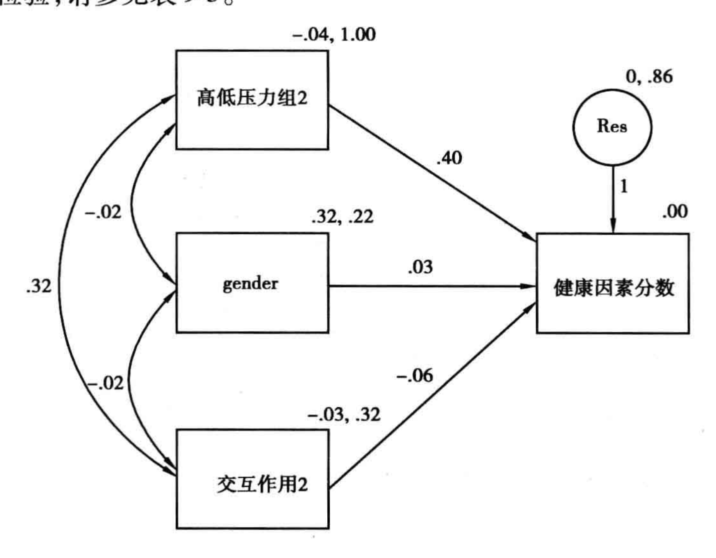


图 9-3 Amos 的路径图设计与分析结果

表 9-3 Amos 路径图中的回归系数报表

			Estimate	S. E.	C. R.	P
健康因素分数	<---	高低压力组 2	.3996	.0631	6.3350	* * *
健康因素分数	<---	gender	.0345	.1123	.3070	.7589
健康因素分数	<---	交互作用 2	-.0645	.1123	-.5744	.5657

由以上分析比较可知，SPSS 与 SEM 这两种不同分析途径，殊途同归，交互作用效果的结果与结论完全相同。

（二）有一个变量为连续变量

以吴丽华(2007)的“工作压力与身心健康”关系研究为例,研究欲探讨两个潜在变量的因果关系会不会因性别而有差异。因身心健康与工作压力均为潜在变量,而性别为一指标变量。因此,需要建立性别与工作压力的交互作用项(交乘积项),较简便的做法是先估计身心健康与工作压力的因素分数,再计算性别(化为虚拟变量)与工作压力的因素分数间的交互作用效果。

图9-4是利用SPSS副程序[Factor Analysis],进行因素分数的计算过程:首先将“工作压力”的指标变量,选入因素分析的变量窗口中,接着打开[Score]的窗口,在打开的因素分数设定窗口中,勾选[Saved as variables],以便将压力因素分数在SPSS Data Editor中建一个新文件(参见图9-5),并在[Method]中勾选一因素分数的计算方法,在本例中,是勾选[Anderson-Rubin]法。“身心健康”的因素分数,亦同法操作。

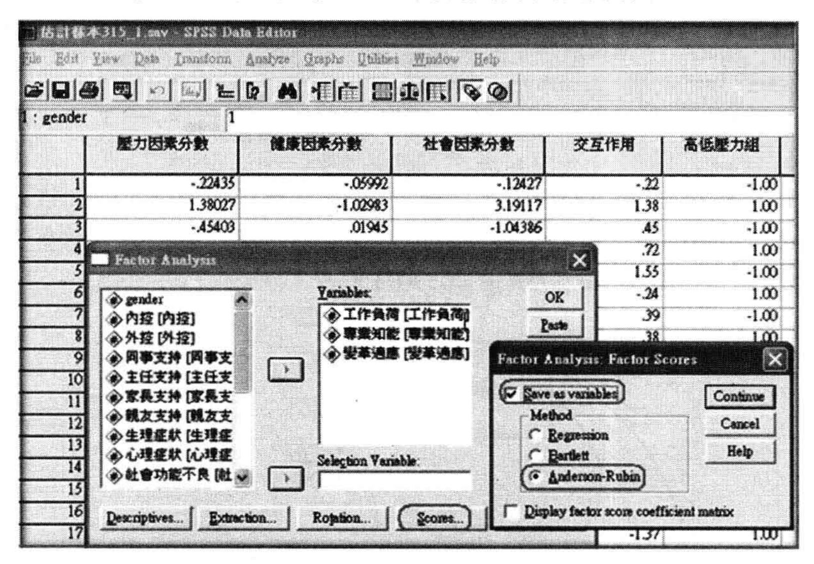


图9-4　SPSS压力因素分数的操作过程

图9-5中的交互作用项分数,是性别(化为虚拟变量)与工作压力分数的交乘积。接着,利用SPSS进行回归分析,所得结果如表9-4,显示出交互作用未达0.05的显著水平(t = -.374, p = 0.709)。本法是利用Dummy coding回归分析法,进行交互作用项的分析。

表9-4　SPSS回归分析报表

		Unstanardized Coefficients		Standardized Coefficients		
Modal		B	Std. Error	Beta.	t	Sig.
1	(Constant)	-.022	.062		-.358	.720
	gendar	.066	.110	.031	.598	.550
	压力因素分数	.440	.063	.440	6.953	.000
	交互作用	-.041	.109	-.024	-.374	.709

数据来源:a. Depend ent Variable Variable:健康因素分数.

估計樣本315-1.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1：壓力因素分數　-0.224354666261175

	gender	壓力因素分數	健康因素分數	社會因素分數	交互作用
1	1	-.22435	-.05992	-.12427	-.22
2	1	1.38027	-1.02983	3.19117	1.38
3	0	-.45403	.01945	-1.04386	.00
4	1	.71778	1.79838	-.20331	.72
5	0	-1.54594	-.26973	.83855	.00
6	0	.24172	.25245	-.37310	.00
7	0	-.38743	-.58446	-.52272	.00
8	1	.37825	-.84556	-1.64299	.38
9	0	-1.36284	.54164	-.72105	.00
10	0	.79436	-1.02747	-.34521	.00

图 9-5　SPSS Data Editor 中的压力因素分数

假如研究者改用 Amos 进行交互作用效果分析，其路径图设计与参数估计值如图 9-6 所示。这些路径参数的统计检验结果如表 9-5，显示出性别与工作压力的交互作用，亦未达 0.05 的显著水平（CR = −0.3758，p = 0.7071）。

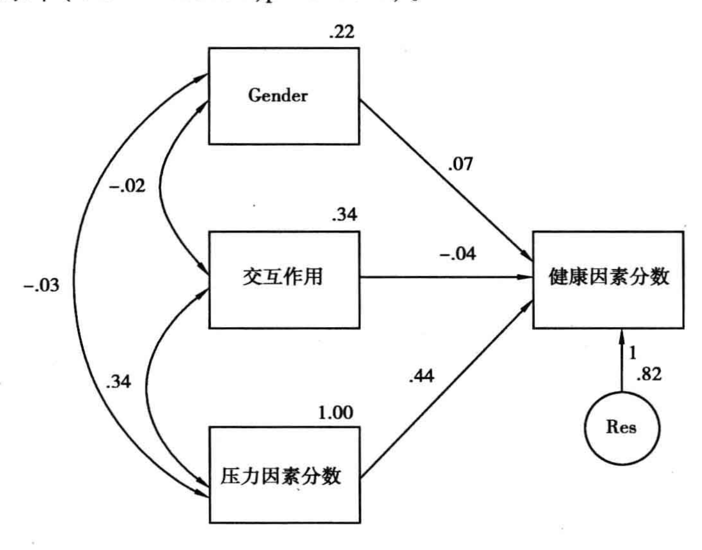


图 9-6　Amos 路径图的设计与相关参数的估计值（交互作用乘积.amw）

表 9-5　Amos 路径图中的回归系数报表

			Estimate	S. E.	C. R.	P
健康因素分数	<---	压力因素分数	.4402	.0630	6.9866	***
健康因素分数	<---	Gender	.0661	.1099	.6011	.5477
健康因素分数	<---	交互作用	−.0406	.1080	−.3758	.7071

由以上分析比较可知,SPSS 与 SEM 两种不同的分析途径,殊途同归,性别与工作压力交互作用效果的结论完全相同。

二、多群组 SEM 法

研究者可以利用因素分数模型或原始分数模型,进行多群组 SEM 分析。首先,研究者需要利用 Amos 的多群组[Manage Group]功能,设定男生与女生两个组别,男生的路径图中路径参数命名为 b1,女生的路径参数命名为 b2,路径图如图 9-7 所示。

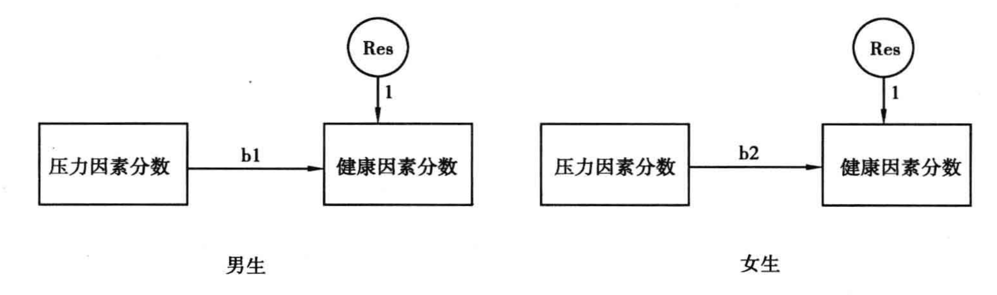


图 9-7　Amos 多群组路径图设计:工作压力与身心健康的关系

多群组设定成功的后,Amos 的主画面就会出现如图 9-8 右侧的男、女生组别的设定。其次,利用 Amos[Manage Model]的功能,在限制模型(Equal slope)中,进行参数等同的设定:b1 = b2,如图 9-9 所示。

Amos 参数的等同设定后,Amos 的主画面就会出现如图 9-8 右侧的 Unconstrained 与 Equal slope 两个理论模型设定。欲观看这两个模型的比较结果,请点开图 9-10 左侧窗口中的[Model Comparison],结果会出现在图 9-10 右侧窗口中。

由图 9-10 所示,所得结果亦发现工作压力与身心健康的关系,不因男女生性别而有差异(χ^2 = 0.1308,df = 1,p = 0.7176)。

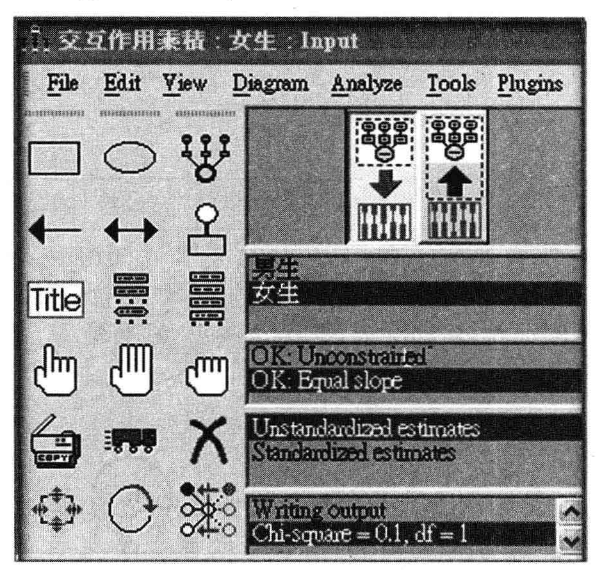


图 9-8　Amos 多群组设定

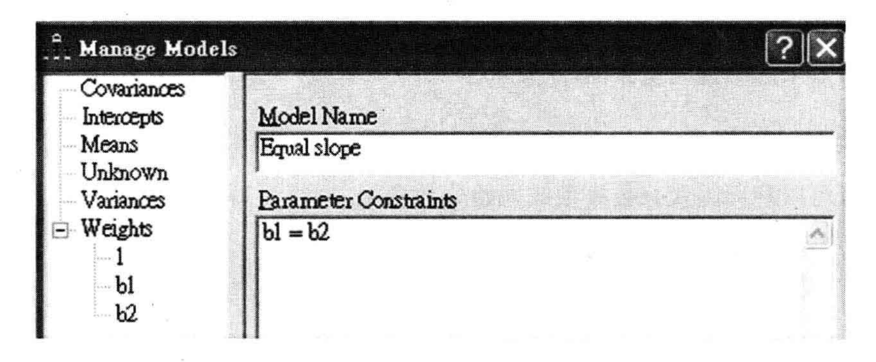
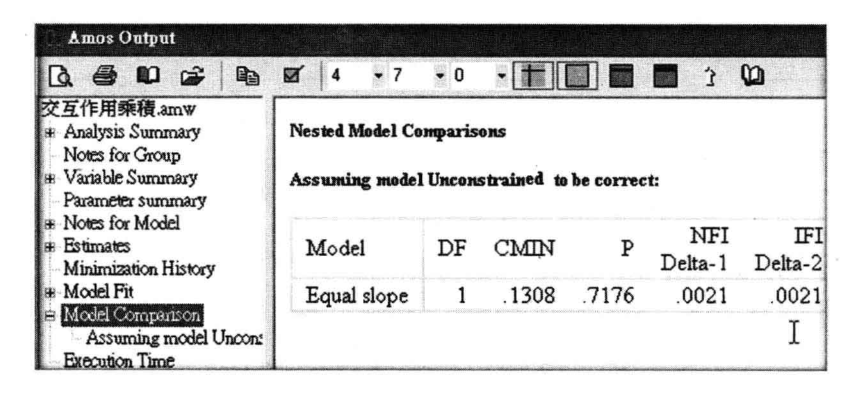


图 9-9　Amos 参数的等同设定



Model	DF	CMIN	P	NFI Delta-1	IFI Delta-2
Equal slope	1	.1308	.7176	.0021	.0021

图 9-10　Amos 嵌套模型比较报表:因素分数模型

研究者如欲利用原始分数,进行多群组 SEM 分析,首先建立如图 9-11 的原始指标与潜在变量路径图,并进行组间路径系数的等同限制(限制过程如同前述),所得的结论亦与前述结果相同(χ^2 = 0.8018,df = 1,p =0.3706),图 9-12 所示。

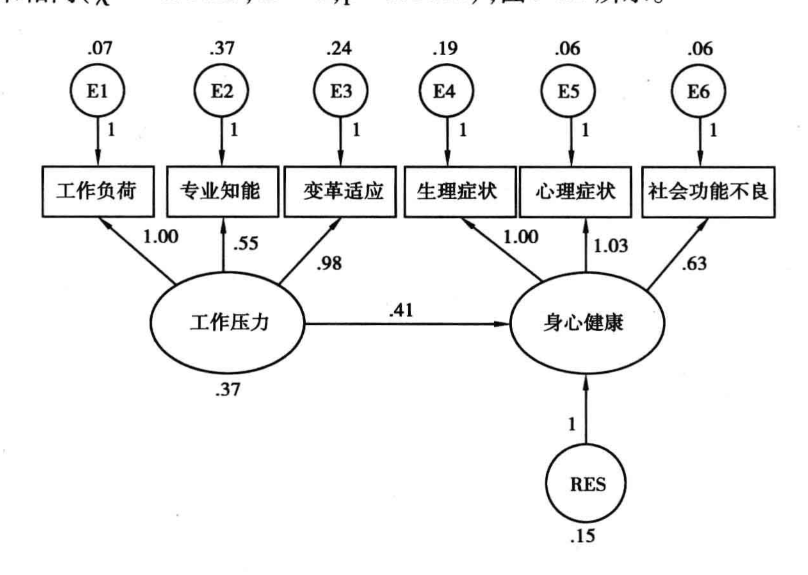
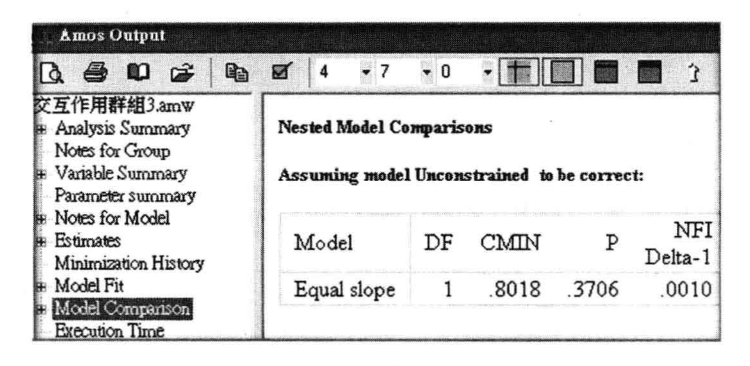


图 9-11　Amos 路径图设计:工作压力与身心健康的关系


Amos Output

交互作用群組3.amw
Analysis Summary
Notes for Group
Variable Summary
Parameter summary
Notes for Model
Estimates
Minimization History
Model Fit
Model Comparison
Execution Time

Nested Model Comparisons

Assuming model Unconstrained to be correct:

Model	DF	CMIN	P	NFI Delta-1
Equal slope	1	.8018	.3706	.0010

图 9-12　Amos 嵌套模型比较报表:原始分数模型

三、因素分量

相比 Ping 的双步骤法,Jonsson's 双步骤因素分量比较简便,首先评估测量模型及计算两个潜在变量下各受试者的因素分数,接着计算各受试者在两个潜在变量下的因素分数乘积的交互作用因子的分数,最后进行这三个潜在变量的回归或 SEM 分析。进行双步骤因素分数交互作用效果的检验,须先检验整个测量模型是否有效(旨在检验单向度的基本假设),如果无效那么整个非线性分析模型亦将是无效的。假如测量模型有效,而又无证据显示有交互作用,研究者最好还是使用线性模型(Jonsson, 1998)。以下仍以吴丽华(2007)的工作压力与身心健康关系的研究为例,研究者如欲探讨这三个潜在变量:工作压力、社会支持与身心健康间整体测量模型的适配性,可绘制如图 9-13 的路径图,结果如表 9-6 所示。

工作压力、社会支持与身心健康的整体测量模型的检验结果如表 9-6 所示。

表 9-6　Amos 整体测量模型检验的输出报表

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	33	74.4235	32	.0000	2.3257
Saturated model	65	.0000	0		
Independence model	20	1257.8075	45	.0000	27.9513

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.9408	.9168	.9654	.9508	.9650
Saturated model	1.0000		1.0000		1.0000
Independence model	.0000	.0000	.0000	.0000	.0000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.0650	.0458	.0844	.0955
Independence model	.2930	.2791	.3070	.0000

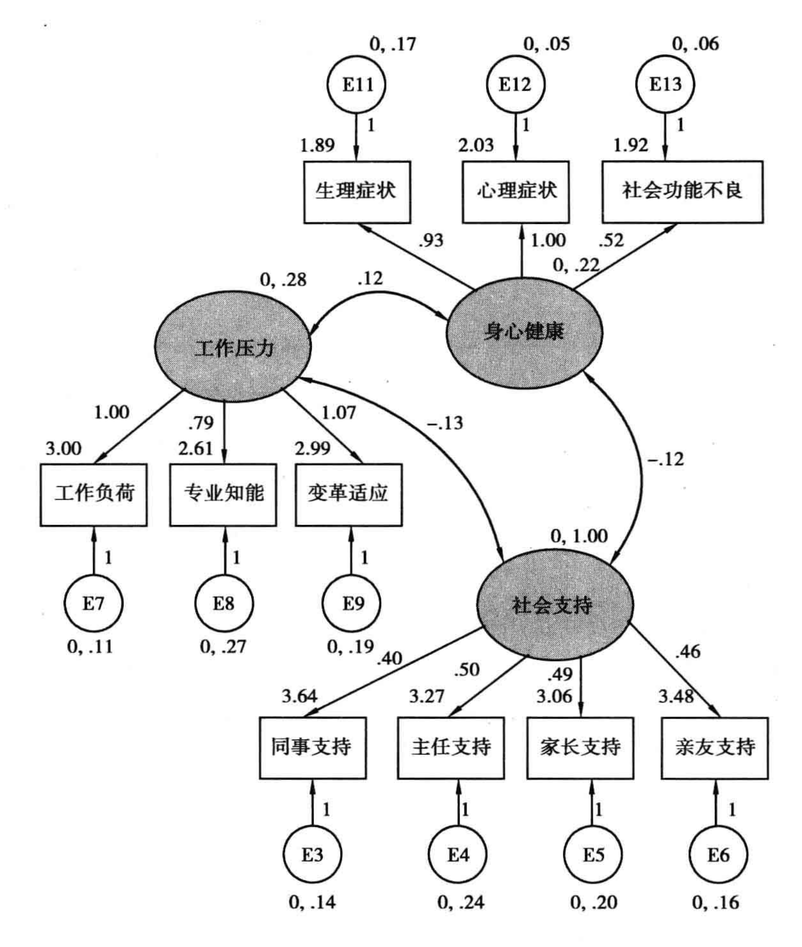


图 9-13　工作压力、社会支持与身心健康的整体测量模型的检验

图 9-13 显示，整体测量模型，虽不尽完美但尚属有效（$\chi^2 = 74.4235$，df = 32，p = 0.0000，CFI = 0.9650，RMSEA = 0.0650）。因此，可继续使用潜在变量的因素分数进行交互作用效果的分析。交互作用项是利用压力因素分数与社会支持因素分数的交乘积，参见图 9-14 中 SPSS 的数据文件的“交互作用 3” 变量。

估計樣本315_1.sav - SPSS Data Editor

File　Edit　View　Data　Transform　Analyze　Graphs　Utilities　Window　Help

1 : gender　1

	壓力因素分數	健康因素分數	社會因素分數	交互作用	高低壓力組	交互作用2	交互作用3
1	-.22435	-.05992	-.12427	-.22	-1.00	-1.00	.03
2	1.38027	-1.02983	3.19117	1.38	1.00	1.00	4.40
3	-.45403	.01945	-1.04386	.45	-1.00	1.00	.47
4	.71778	1.79838	-.20331	.72	1.00	1.00	-.15
5	-1.54594	-.26973	.83855	1.55	-1.00	1.00	-1.30
6	.24172	.25245	-.37310	-.24	1.00	-1.00	-.09
7	-.38743	-.58446	-.52272	.39	-1.00	1.00	.20
8	.37825	-.84556	-1.64299	.38	1.00	1.00	-.62
9	-1.36284	.54164	-.72106	1.36	-1.00	1.00	.98
10	.79436	-1.02747	-.34521	-.79	1.00	-1.00	-.27
11	2.56539	-.13909	.32422	-2.57	1.00	-1.00	.83

图 9-14　SPSS 数据文件：工作压力、社会支持与身心健康的因素分数

由图9-15与表9-7中Amos的分析结果知,交互作用未达0.05显著水平(CR = 1.1676,p = 0.2430);研究者如利用SPSS进行回归分析,其结果如表9-8所示,交互作用亦未达0.05显著水平(t = 1.162,p = 0.246),所得结论亦完全相同。

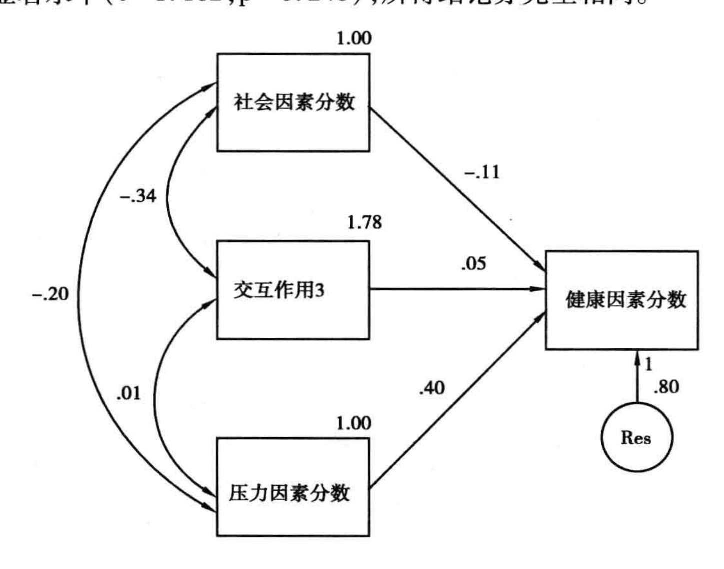


图9-15　Amos工作压力与社会支持交互作用分析结果:因素分数模型

表9-7　Amos因素分数模型的输出报表

			Estimate	S. E.	C. R.	P
健康因素分数	<---	压力因素分数	.4021	.0517	7.7804	* * *
健康因素分数	<---	社会因素分数	-.1053	.0534	-1.9716	.0487
健康因素分数	<---	交互作用3	.0457	.0392	1.1676	.2430

表9-8　SPSS交互作用的回归分析结果:因素分数模型

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta.	t	Sig.
1(Constant)	.009	.051		.181	.857
压力因素分数	.402	.052	.402	7.743	.000
社会因素分数	-.105	.054	-.105	-1.962	.051
交互作用3	.046	.039	.061	1.162	.246

a. Dependent Variable Variable:健康因素分数.

四、Ping的单指标双步骤法

在本法的示范中,仍沿用吴丽华(2007)的研究数据,参见图9-16中的变量与相关数据,前面四个变量(同事支持～亲友支持)是社会支持因子,接着3个变量(工作负～变革适应)是压力因子,其次三个变量(生理症状～社会功能不良)是健康因子,最后则为交互作用项。Ping的单指标双步骤法,在计算交互作用项前,除了交互作用项外,最好将所有

自变量与依变量化为离均差分数后(参见图 9-16),再将各潜在变量下所属的指标加总平均作为一新变量,接着利用两个加总平均变量的积作为交互作用项。以第一个受试者为例,其交互作用的分数等于 0.01435 = [((0.36 - 0.27 - 0.86 + 0.36)/4) * ((0.17 + 0.06 - 0.65)/3)]。前述的数据转换及交互作用项的建立是 Ping 的单指标双步骤法的先导工作。

接着,研究者须评估测量模型的适配性及潜在变量的向度(dimensionality),模型适配及单向度的假设符合后,根据测量模式所得到的参数估计值,计算交互作用项的因素负荷量与误差变异量,这是 Ping 的单指标双步骤法在第一步骤必须完成的工作。根据模型适配及单向度的评估指标,图 9-18 底部 X(代表工作压力)、Z(代表社会支持)与 XZ 的信度(reliability)分别为 0.7979,0.8232 与 0.6755,而 Amos 测量模型的适配度指标(参见图 9-17 中的 GFI = 0.952,CFI = 0.947,RMSEA = 0.102),反映出模型的适配性及单向度假设,尚差强人意。

估計樣本315-mean-centering.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

	同事支持	主任支持	家長支持	親友支持	工作負荷	專業知能	變革適應	生理症狀	心理症狀	社會功能不良	平均交互
1	.36	-.27	-.86	.36	.17	.06	-.65	.11	-.03	-.09	.01
2	1.36	1.73	1.94	1.52	.50	1.39	.51	-.89	-.23	-.26	1.31
3	-.04	-.87	-1.06	-.31	-.33	-.61	.18	-.06	-.03	.08	.14
4	-.44	.53	-.06	-.31	.33	.73	.18	.94	.97	.41	-.03
5	.16	.53	.54	.52	-1.00	-.61	-.99	-.22	-.03	-.09	-.38
6	-.44	-.47	.14	.02	.17	.06	.18	.11	.37	-.09	-.02

图 9-16　SPSS 数据文件:工作压力、社会支持与身心健康的离均差分数

假如单一构念的测量模型不适配的话,那么利用 X 与 Z 在测量模型中的估计值,所间接计算出来的 XZ 因素负荷量与误差变异量,将与包含 XZ 交互作用项的全模型中的估计值具有差异,而导致偏估,结构模型的适配度亦会因此下降。万一 X 和 Z 在测量模型与结构模型中的估计值不相同(例如,第一或第二个小数位应相等),研究者应利用全模型中的估计值重新再计算交互作用项的因素负荷量与误差变异量,并重新评估 X,Z 在测量模型与结构模型中的估计值的等同性。如此重新分析一、二次,应能达到精确的要求。请比较图 9-17 与图 9-20 中,工作压力(X)及社会支持(Z)的路径系数的等同性。

因为 Amos 目前无法直接进行交互作用项的参数估计,所以如同前述多元指标法,研究者必须先利用第一阶段测量模型(如图 9-17 所示)的估计值,换算这个非线性指标的因素负荷量与误差变异量,以作为第二阶段结构模型中的固定常数。换算这个交互作用项的因素负荷量与误差变异量,在正态性与单向度的基本假设下,所需的公式描述如下:

1. 交互作用项的因素负荷量为

$$\lambda_{x:z} = \Lambda_X \Lambda_Z, \text{式中 } \Lambda_X = (\lambda_{x1} + \lambda_{x2} + \lambda_{x3})/3,$$

$$\theta_X = (Var(\varepsilon_{x1}) + Var(\varepsilon_{x2}) + Var(\varepsilon_{x3}))/3,$$

$$\Lambda_Z = (\lambda_{z1} + \lambda_{z2} + \lambda_{z3} + \lambda_{z4})/4,$$

$$\theta_Z = (Var(\varepsilon_{z1}) + Var(\varepsilon_{z2}) + Var(\varepsilon_{z3}) + Var(\varepsilon_{z4}))/4。$$

2. 交互作用项的误差变异量为

$$\theta\varepsilon_{x:z} = \Lambda\chi^2 \mathrm{Var}(X)\theta_Z + \Lambda_Z^2 \mathrm{Var}(Z)\theta_x + \theta_x\theta_Z。$$

以交互作用项的因素负荷量为例，在工作压力上 $\Lambda_X = (1.0 + 0.77 + 1.04)/3 = 0.9367$，在社会支持上 $\Lambda_Z = (1.0 + 1.25 + 1.23 + 1.15)/4 = 1.160$；带入公式 $\lambda_{x:z} = 0.9367 * 1.160 = 1.08653$。可见计算过程稍嫌繁琐，研究者可使用 Ping's 交互作用项参数的 Excel 计算器(Ping, 1995)，参见图 9-18。这一阶段性的任务达成后，就可进入第二阶段的 SEM 分析。

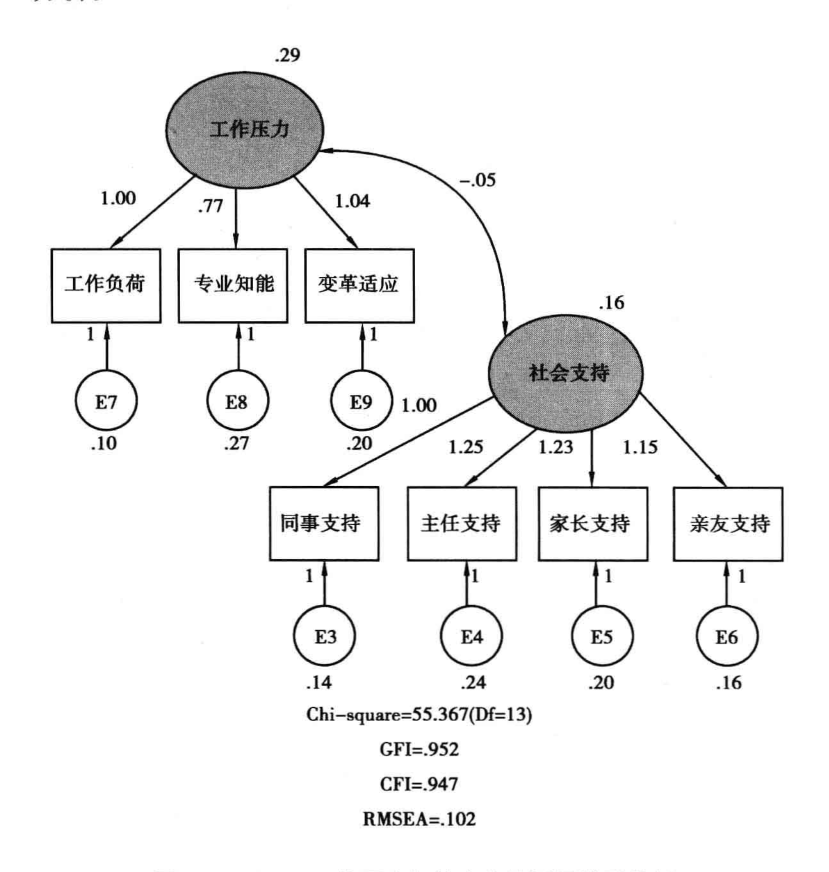
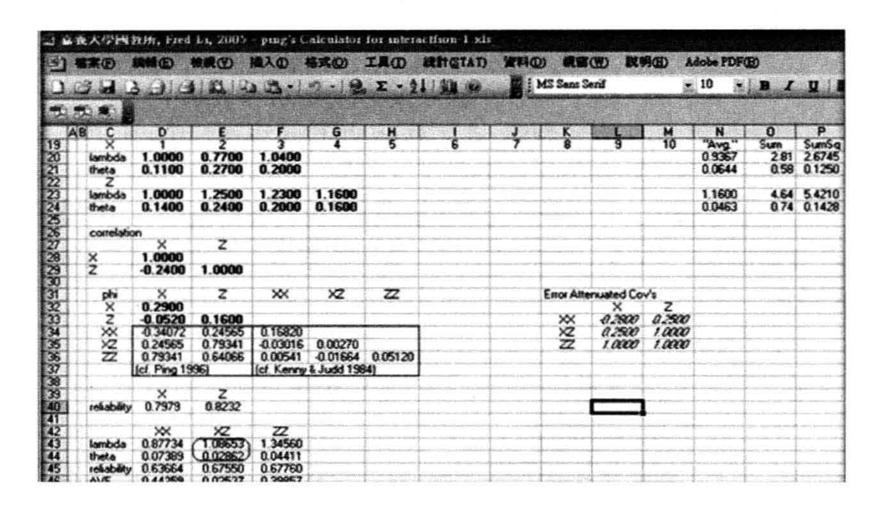


图 9-17　Amos 工作压力与社会支持测量模型分析

由图 9-18 知，交互作用项的因素负荷量与误差变异量，分别为 1.08653 与 0.02862，将这两个参数估计值带入图 9-19 相对应的路径中作为常数（参见交互作用变量的标准路径系数与误差变异量），再估计交互作用项与身心健康关系的效果值。

为了模型适配度与交互作用项的相关参数能获得较稳定的结果，交互作用与工作压力、社会支持间的共变关系在结构模型中亦应释放进行估计，参见图 9-19 的路径关系。

由图 9-20 与表 9-9 知，工作压力与社会支持对身心健康的交互作用效果(0.0791)未达统计上的 0.05 显著水平（CR = 0.8223，p = 0.4109）。Ping's 单指标双步骤模型分析所得结果与前述 Jonsson's 双步骤因素分数模型分析的结论相同。



	C	D	E	F	G	H	I	J	K	L	M	N	O	P
19	X	1	2	3	4	5	6	7	8	9	10	"Avg."	Sum	SumSq
20	lambda	1.0000	0.7700	1.0400								0.9367	2.81	2.6745
21	theta	0.1100	0.2700	0.2000								0.0644	0.58	0.1250
22	Z													
23	lambda	1.0000	1.2500	1.2300	1.1600							1.1600	4.64	5.4210
24	theta	0.1400	0.2400	0.2000	0.1600							0.0463	0.74	0.1428
25														
26	correlation													
27		X	Z											
28	X	1.0000												
29	Z	-0.2400	1.0000											
30														
31	phi	X	Z	XX	XZ	ZZ			Error Attenuated Cov's					
32	X	0.2900								X	Z			
33	Z	-0.0520	0.1600						XX	-0.2800	0.2500			
34	XX	-0.34072	0.24565	0.16820					XZ	0.2500	1.0000			
35	XZ	0.24565	0.79341	-0.03016	0.00270				ZZ	1.0000	1.0000			
36	ZZ	0.79341	0.64066	0.00541	-0.01664	0.05120								
37		(cf. Ping 1996)		(cf. Kenny & Judd 1984)										
38														
39		X	Z											
40	reliability	0.7979	0.8232											
41														
42		XX	XZ	ZZ										
43	lambda	0.87734	1.08653	1.34560										
44	theta	0.07389	0.02862	0.04411										
45	reliability	0.63664	0.67550	0.67760										

图 9-18　Ping's 交互作用项参数的 Excel 计算器

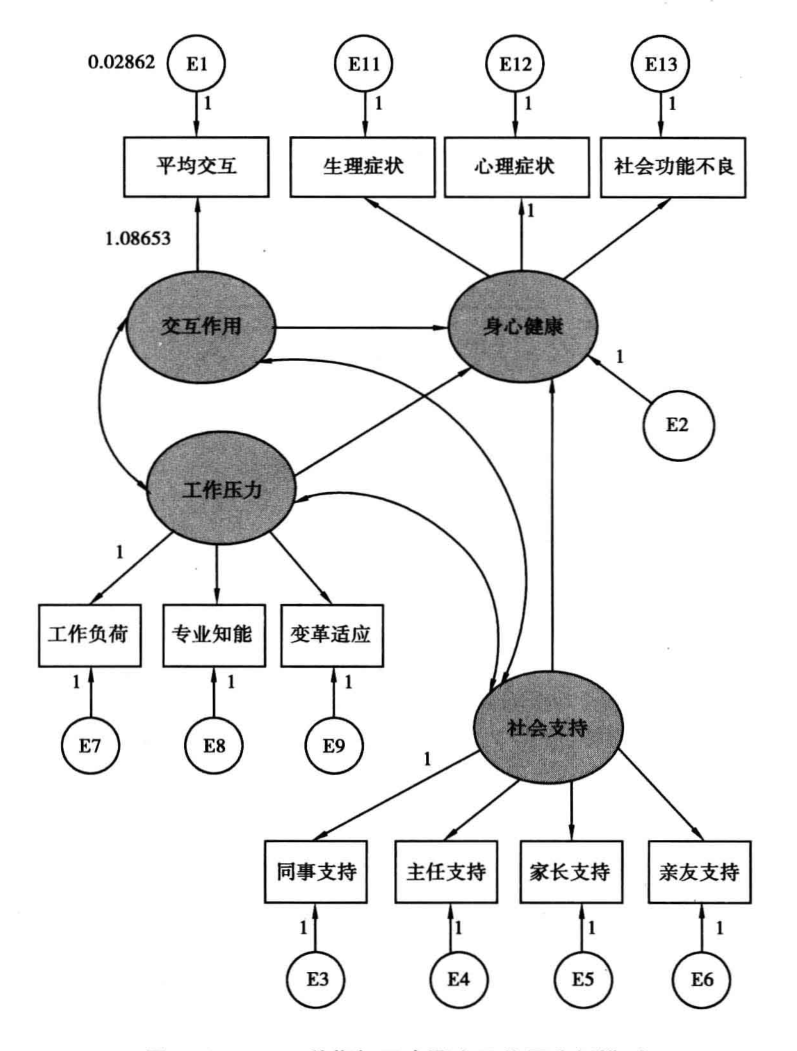


图 9-19　Ping's 单指标双步骤交互作用分析模型

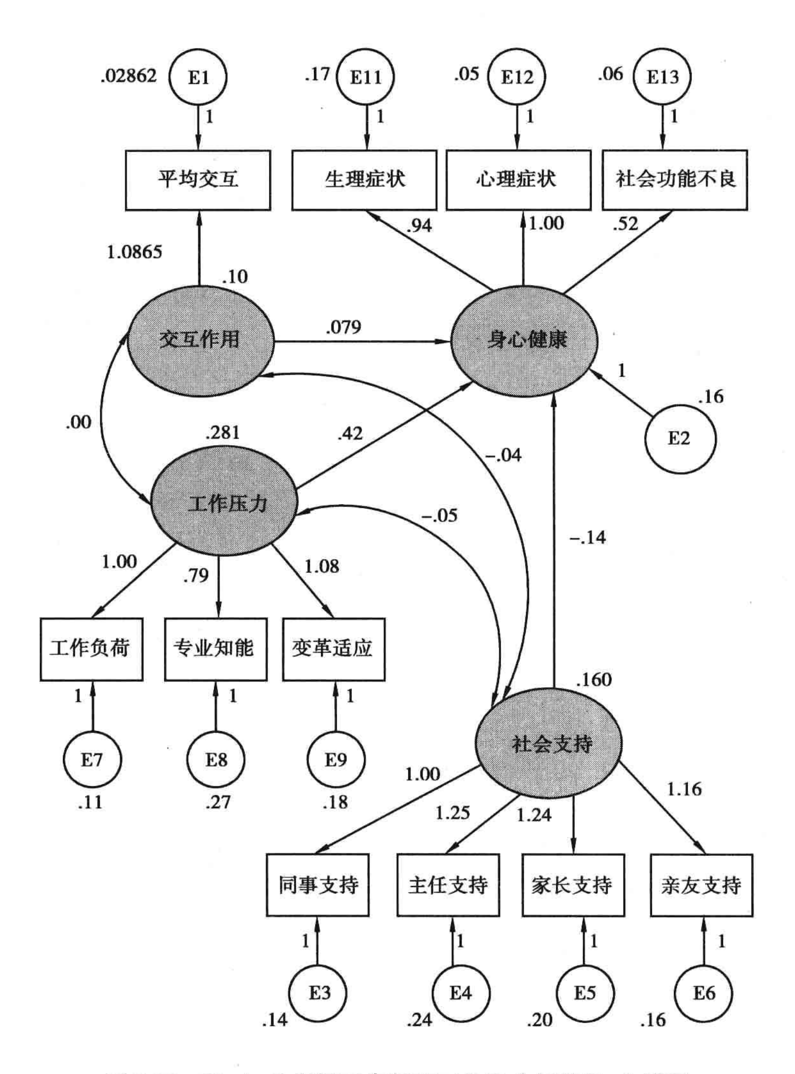


图 9-20　Ping's 单指标双步骤交互作用分析结果:全模型

表 9-9　Amos 路径系数摘要表

			Estimate	S. E.	C. R.	P
身心健康	<---	社会支持	-.1402	.0794	-1.7654	.0775
身心健康	<---	工作压力	.4160	.0592	7.0261	* * *
身心健康	<---	交互作用	.0791	.0962	.8223	.4109
生理症状	<---	身心健康	.9366	.0727	12.8784	* * *
社会功能不良	<---	身心健康	.5202	.0410	12.6881	* * *
变革适应	<---	工作压力	1.0772	.0852	12.6486	* * *
主任支持	<---	社会支持	1.2514	.1117	11.2068	* * *
专业知能	<---	工作压力	.7937	.0746	10.6395	* * *
亲友支持	<---	社会支持	1.1592	.0986	11.7582	* * *
家长支持	<---	社会支持	1.2397	.1076	11.5201	* * *

第二节 直接效果的检验方法

同样地,直接效果的检验亦可以通过回归分析与 SEM 分析,以下仍以吴丽华(2007)的社会支持、教师工作压力与身心健康关系的 SEM 提议模型为例。为了检验它们三者间的直接效果,在路径图的设计阶段,利用 Amos 的对象属性窗口,特以 a 标注教师工作压力与身心健康间的关系,以 b 标注社会支持与教师工作压力间的关系,以 c 标注社会支持与身心健康间的关系,命名结果如图 9-21 所示。

利用 Amos 执行统计分析后,社会支持、教师工作压力与身心健康间的直接效果与相关的统计检验结果,显示于表 9-10 中。由表中的相关 p 值知,路径系数 a, b 与 c 的直接效果,均达统计上 0.05 的显著水平($p < 0.05$),拒绝了路径系数为 0 的虚无假设,亦即它们之间均具有某种程度上的影响力。在该研究中,"社会支持量表"总分越高,代表教师社会支持越大;而"身心健康量表"总分越低,代表教师身心健康越佳;"教师工作压力量表"总分越高,代表教师工作压力越大。因此,根据统计检验结果,可获得以下结论:①教师社会支持越大时,教师工作压力就越小;②教师工作压力越大时,教师身心健康就越差;③教师社会支持越大时,教师身心健康就越佳。

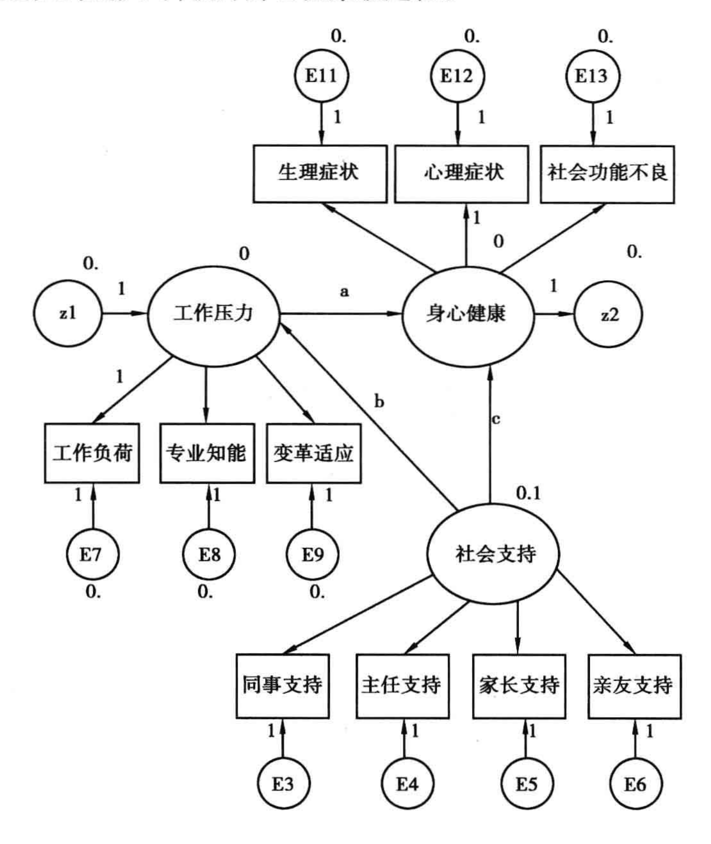


图 9-21 社会支持、教师工作压力与身心健康关系的提议模型

表 9-10　Amos 路径系数的输出报表：直接效果

Regression Weights.（Group number 1-Default model）

			Estimate	S. E.	C. R.	P	Label
工作压力	<---	社会支持	-.1302	.0360	-3.6164	* * *	b
身心健康	<---	工作压力	.4114	.0588	6.9976	* * *	a
身心健康	<---	社会支持	-.0648	.0294	-2.2026	.0276	c

第三节　间接效果检验的基本假设与流程

研究者在分析间接或中介效果值时，首先需要评估一些基本假设是否成立。例如：①中介变量与结果变量具有因果关系；②中介变量没有测量误差；③没有遗漏任何中介变量（M）与结果变量（Y）的致因变量（variables that cause M and Y）；④预测变量与中介变量无交互作用（MacKinnon，Fairchild，Fritz，2007）。上述这些假设可能不容易检验，但研究者可提供理论依据或相关文献予以支撑。

至于间接效果的检验流程，采用最广的、最经典的首推 Baron 与 Kenny（1986）所提的四个步骤：

1. X 与 Y 的零阶相关必须具有显著相关；
2. X 与中介变量零阶相关必须具有显著相关；
3. 中介变量与 Y 的净相关必须具有显著相关（排除 X 对 Y 的直接关系）；
4. 判断中介效果是完全中介或部分中介。这四个步骤的细节，请参见第四节因果法，在此不赘述。

这四个步骤也遭受过一些批评，例如：第一步骤似乎不是必要的。因此，Shrout 与 Bolger（2002）就建议利用图 9-22 的流程进行间接或中介效果值的分析。首先，研究者须判断中介变量与起始变量（X）或结果变量（Y）的关系，到底何者较密切？如果中介变量与起始变量（X）关系较密切，属近程中介效果（路径 a 系数属于 proximal effect，路径 b 系数势必相对较小），研究者就须进行路径 c 的估计与检验；如果中介变量（M）与结果变量（Y）关系较密切，属远程中介效果（属于 distal effect，路径 c^* 系数势必会较小），研究者就可跳过第一步总效果的检验，直接跳到第二步的统计检验。另外，当中介变量具有抑制变量的特质时，研究者会发现总效果（直接效果 + 间接效果）接近或等于 0。因此，当中介变量具有抑制变量的现象时，虽然总效果值不大（可能由于直接效果与间接效果的参数估计值符号相反，亦即关系强度相若，但方向相反），但中介效果可能是存在的，因而可以跳过第一步的检验，直接跳到第二步的统计检验。最后，Shrout 与 Bolger（2002）建议使用中介效果与总效果的比值，作为中介效果强度的评估指标。

另外，研究者如欲获得中介效果与直接、间接效果的置信区间，可以利用 Amos [Analysis properties] 下的 [Output] 按钮，点开后勾选 [Indirect，direct and total effects]，接着，勾选 [Perform bootstrap] 与 [Percentile confidence interval]，并设定样本复制的个数，研究者即可获得中介效果与直接、间接效果的置信区间。

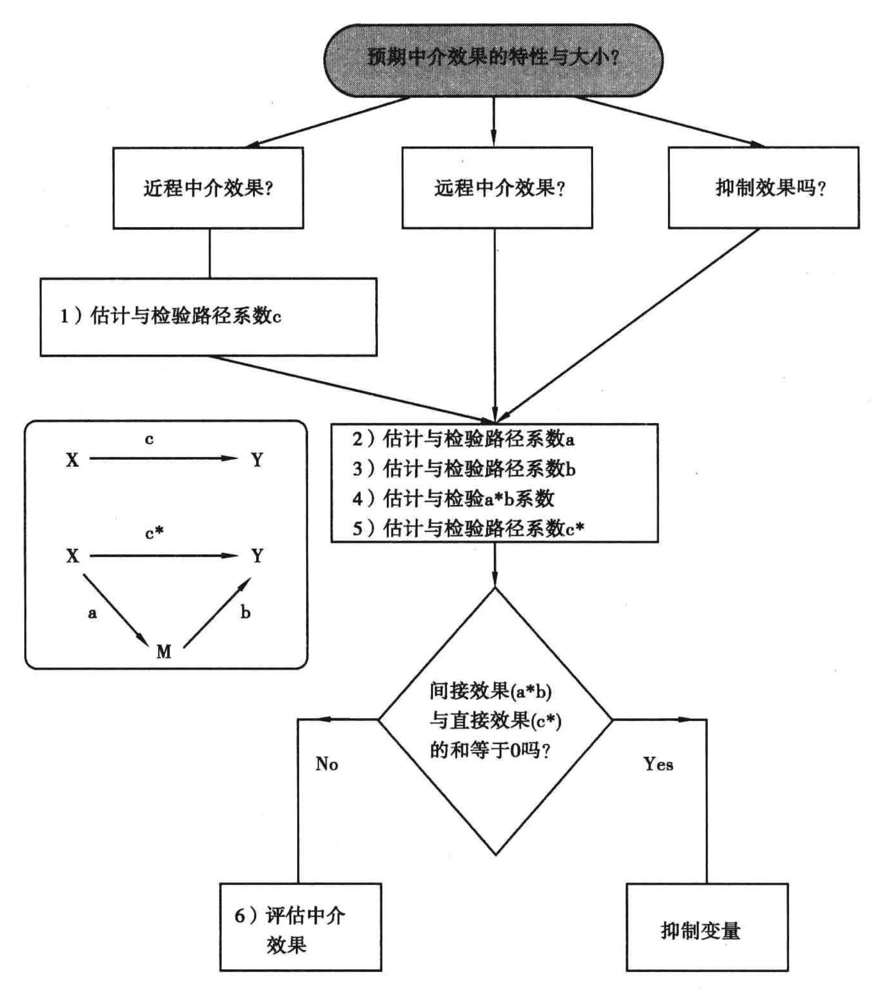


图 9-22　间接效果检验的流程

注:修订定自图 8,Shrout & Bolger(2002)

第四节　间接效果的检验方法

综合 MacKinnon 与 Lockwood 等人(2002),Shrout 与 Bolger(2002),James、Mulaik 与 Brett(2006)与 Cheung(2007)的归类,中介变量效果的统计检验方法,可细分为四类:因果法、系数差异法、系数乘积法与 Bootstrap 法,兹逐一介绍于后:

一、因果法

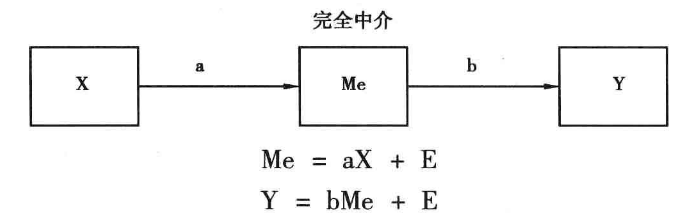
因果检验法又称 Baron 与 Kenny 法，事实上是一种回归分析法，根据 Baron 与 Kenny (1986)的主张,欲证明间接效果,须进行以下四步骤的分析:

1. X 与 Y 的零阶相关必须具有显著相关,亦即回归方程:$Y = \beta_{10} + \beta_{11}X + \varepsilon_1$,式中的整体处理效果(X)的系数 β_{11},须达显著水平。
2. X 与中介变量 Me 零阶相关必须具有显著相关,亦即回归方程中 $Me = \beta_{20} +$

$\beta_{21}X + \varepsilon_2$，X 对于中介变量效果的系数 β_{21}，须达显著水平。

3. 中介变量 Me 与 Y 的净相关必须具有显著相关（排除 X 对 Y 的直接关系），亦即回归方程中 $Y = \beta_{30} + \beta_{31}X + \beta_{32}Me + \varepsilon_3$，其中介变量（Me）对于依变量效果的净回归系数 β_{32}（排除 X 的影响力），须达显著水平，否则 Me 与 Y 间的一阶相关即为虚假相关（spurious relationship）。
4. 当排除中介变量 Me 对 Y 的直接关系后，X 即不再与 Y 具有任何显著相关（$\beta_{31} = 0$），我们称的为完全中介；有部分中介存在时，X 与 Y 的净相关会小于 X 与 Y 的零阶相关（$\beta_{31} < \beta_{11}$）。换言之，间接效果（$\beta_{31} - \beta_{11}$）会等于 β_{21} 与 β_{32} 的乘积。

在此特别指出 Baron 与 Kenny（1986）的间接效果检验方法与目前流行的 SEM 分析法的差异处。当研究者想利用 SEM 分析而提出完全中介效果的模型时，其所涉及的路径图与回归方程为：

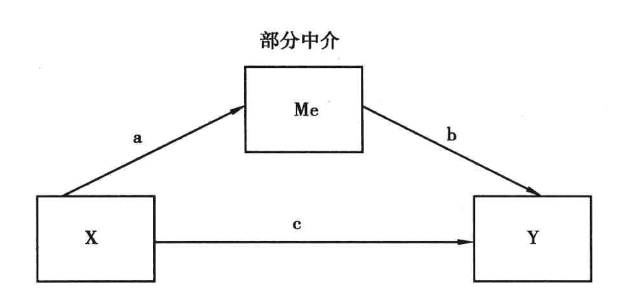


$$Me = aX + E$$

$$Y = bMe + E$$

式中，a，b 两个回归系数，是分别由上述两个独立回归方程计算而得。此完全中介模型，假设总效果不为 0，但直接效果等于 0，因此 X 与 Y 独立无关。

当研究者想利用 SEM 分析而提出部分中介效果的模型时，须使用以下的路径图与回归方程：



$$Me = aX + E$$

式中，a 是上述回归方程的回归系数。

$$Y = cX + bMe + E$$

式中，c，b 两个净回归系数，由上述第二个回归方程计算而得，c 是排除 Me，b 是排除 X 后的净回归系数。由此观之，研究者所提议的模型设定：部分中介效果模型或完全中介效果模型，会影响到相关回归系数的计算，也影响到中介效果的计算。如果研究者是利用 Baron 与 Kenny（1986）的中介效果检验模型，则不分“部分中介”或“完全中介”，一律假设为部分中介效果。因此，中介效果的计算方法与 SEM 中的相同（James，Mulaik，& Brett，2006）。不过，当 X 是抑制变量（suppressor variable）时，即使 X 与 Y 没有关联性（总效果等于 0 或接近于 0），亦有可能存在中介效果（例如，间接效果与直接效果的大小相同，但关系方向相反或完全中介时）。因此，MacKinnon，Lockwood，Hoffman，West 与 Sheets（2002）认为本法统计检验力较弱。这是因果法运用上的盲点，应用者应谨慎运用它。

二、系数差异法

本法旨在计算 X 与 Y 的零相关以及 X 与 Y 的净相关(排除 Me 对于 Y 的效果):$r_{x,y} - r_{x,y.Me}$(以下简称 r-r′)。此相关系数差异的统计检验方法有很多种,在此仅列出 Freedman 与 Schatzkin(1992)的计算公式:

$t_{N-2} = \frac{r-r'}{\sigma_{r-r'}}$,式中,$\sigma_{r-r'} = \sqrt{\sigma_r^2 + \sigma_{r'}^2 - 2\sigma_r\sigma_{r'}\sqrt{1-\rho_{XI}^2}}$($\rho_{XI}$表示自变量与中介变量的相关系数,$\delta_r$表示 r 系数的标准误,$\delta_{r'}$表示 r′ 系数的标准误。所得的统计量可以利用 t 分布($df = N-2$)进行显著性检验。

即使中介变量 I 与 Y 完全没有关联性,因净相关全距的限制,亦有可能存在中介效果。这是本法运用上的限制,应用者应谨慎运用它(MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002)。

三、系数乘积法

通过系数乘积计算间接效果的检验方法, 主要有 Z 检验法、幽灵变量(Phantom variable)法与 Bayesian SEM 自订参量。兹分别简介如下:

(一)Z 检验法

Z 检验的公式:$z = \frac{ab}{SE_{ab}}$,亦即 z 统计量为间接效果除以其标准误,用以检验间接效果系数的总体为 0(即间接效果的虚无假设 $H_o: \alpha\beta = 0$)。本法常用的公式有三:

* Sobel 检验公式

$$z - value = \frac{a \times b}{\sqrt{b^2 \times s_a^2 + a^2 \times s_b^2}}$$

* Aroian 检验公式

$$z - value = \frac{a \times b}{\sqrt{b^2 \times s_a^2 + a^2 \times s_b^2 + s_a^2 + s_b^2}}$$

* Goodman 检验公式

$$z - value = \frac{a \times b}{\sqrt{b^2 \times s_a^2 + a^2 \times s_b^2 - s_a^2 + s_b^2}}$$

其中以 Sobel(1982)的 z 检验法最常用,S_a表示回归系数 a 的标准误,S_b表示回归系数 b 的标准误。运用本法进行 z 检验,应用者应考虑这些统计量在样本不大时,可能并非标准正态分布,从而导致统计检验力不足(Mackinnon et al. ,1998; Cheung, 2007)。

上述三种统计量的计算,可利用 Preacher 与 Leonardelli(2006)的网络计算器,网址为 http://www.psych.ku.edu/preacher/sobel/sobel.htm. 附录三的 Amos VB. NET 程序,亦可参考使用。

研究者只要输入 a 与 b 及相关的标准误,就可立即获得三种检验结果,甚为便利。实例解说请参见本章第六节的说明。

(二)幽灵变量法

幽灵变量法(Rindskopf, 1984; Cheung, 2007),就是将一个没有实质意义的变量加

入模型中，旨在进行非线性参数的限制。研究者可以利用 LISREL 的程序特性，建立一个交互作用的幽灵变量（Phantom variable），以估计其置信区间并进行相关的统计检验，这是 LISREL 独特的参数限制功能。目前 Amos 尚未提供此幽灵变量的设计，但其 Bayesian SEM 则有类似功能。因此，以下只提供 LISREL 的程序语法与输出的报表，并依中介变量的个数与因果关系：单一中介变量、关联双中介变量与无关联双中介变量等三类中介形式，逐一利用实际数据说明如下。

	Input:		Test statistic:	p-value:
a		Sobel test:		
b		Aroian test:		
s_a		Goodman test:		
s_b		Reset all	Calculate	

1. 单一中介变量

幽灵变量法，研究者须建立一个幽灵变量，作为独立变量（如同事支持）的预测变量，并将其平均数与标准差设定为 0，参见图 9-23 的路径设计。因此，幽灵变量加入模型并不会影响模型与参数的估计。本路径设计 LISREL 会自动估计间接效果 a * b 参数。

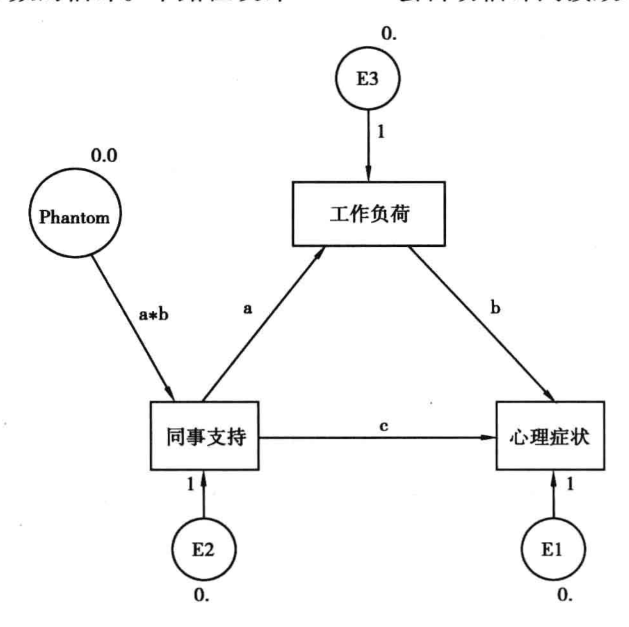


图 9-23　LISREL 幽灵变量的路径设计

兹以吴丽华（2007）的研究数据为例，LISREL 的程序语法与输出报表，逐一条列于后。使用者如对 LISREL 程序的操作方法有疑问，请参考相关书籍，不在此赘述。

※LISREL 程序设计

```
TI A Phantom Var is created to test the indirect effect
DA NI = 3 NO = 315 NG = 1 MA = CM AP = 1
```

```
! Create one additional free parameter

LA
同事支持   工作负荷   心理症状

! Covariance matrix

CM
0.30
-0.03  0.40
-0.04  0.12  0.27

! Means

ME
3.64  3.00  2.04

! 选取变量顺序,先 Y 后 X

SE
2  3  1/

! 模型规划

MO NX = 1 NY = 2 BE = FU, FI GA = FU, FI PH = SY, FR PS = DI, FR TY = FI TX = FI
AL = FI KA = FI
FR BE(2, 1) GA(1, 1) GA(2, 1) AL(1) AL(2) KA(1)

! Create an interaction term

CO PAR(1) = GA(1, 1) * BE(2, 1)
PD
OU ME = ML IT = 250   ND = 3
```

上述 LISREL 程序中,CO PAR(1)是用来建立一个新的待检验参数:GA(1,1) * BE(2, 1),即中介效果。

※Lisrel 输出结果

TI A Phantom Var is created to test the indirect effect

Covariance Matrix

	工作负荷	心理症状	同事支持
工作负荷	0.400		
心理症状	0.120	0.270	
同事支持	-0.030	-0.040	0.300

Means

工作负荷	心理症状	同事支持
3.000	2.040	3.640

Parameter Specifications

BETA

	工作负荷	心理症状
工作负荷	0	0
心理症状	1	0

```
          GAMMA
                 同事支持
工作负荷            2
心理症状            3
          PHI
                 同事支持
                    4
          PSI
                 工作负荷        心理症状
                    5              6
          ALPHA
                 工作负荷        心理症状
                    7              8
LISREL Estimates (Maximum Likelihood)
          BETA
                 工作负荷        心理症状
工作负荷           - -             - -
心理症状          0.292            - -
                (0.043)
                  6.780
          GAMMA
                 同事支持
工作负荷          -0.100
                 (0.065)
                 -1.538
心理症状          -0.104
                 (0.050)
                 -2.092
          Covariance Matrix of Y and X
                 工作负荷        心理症状        同事支持
工作负荷          0.400
心理症状          0.120           0.270
同事支持         -0.030          -0.040           0.300
          Mean Vector of Eta-Variables
                 工作负荷        心理症状
                  3.000           2.040
PHI
                 同事支持
                  0.300
                 (0.024)
                  12.510
PSI
注:This matrix is diagonal.
                 工作负荷        心理症状
                  0.397           0.231
                 (0.032)         (0.018)
                  12.510          12.510
Squared Multiple Correlations for Structural Equations
                 工作负荷        心理症状
                  0.007           0.145
Squared Multiple Correlations for Reduced Form
                 工作负荷        心理症状
                  0.007           0.020
```

```
        Reduced Form
                同事支持
工作负荷          -0.100
                (0.065)
                 -1.538
心理症状          -0.133
                (0.053)
                 -2.511
        ALPHA
                工作负荷      心理症状
                  3.364       1.542
                (0.239)     (0.233)
                 14.055       6.618
        ADDITIONAL PARAMETERS
                PA(1)
                -0.029
                (0.019)
                -1.500
The model is saturated, the fit is perfect!
```

LISREL 报表中 Beta 与 Gamma 矩阵中的数据显示,中介效果的参数为 -0.029 (= -0.1 * 0.292),其 t 检验值为 -1.50(= -0.029/0.019),未达统计上的 0.05 显著水平,亦即"工作负荷"变量没有显著的中介效果,研究者亦可使用该参数的标准误 0.019,建立 0.05 置信区间:-0.029 ± 1.96 * 0.019。

2. 双中介变量(中介变量间具有因果关系)

这是单一中介变量分析的延伸,LISREL 程序的撰写非常类似,图 9-24 中的"内控"与"工作负荷"皆为中介变量,而且这两个中介变量间具有因果关系。本路径设计 LISREL 会自动估计间接效果 A * B * C 参数。

兹以吴丽华(2007)的研究数据为例,LISREL 的程序语法与输出报表,依序条列于后:

※LISREL 程序设计

```
The following lines were read from file F:\lisrel852\Indirect2.lpj:
TI To create a phantom var for testing indirect effect:2 mediators
DA NI = 4 NO = 315 NG = 1 MA = CM AP = 1
LA
内控  同事支持  工作负荷  心理症状
CM
.271
 -.061  .301
.061  -.029  .396
.067  -.036  .116  .270
ME
2.3833  3.6382  3.0011  2.0349
SE
内控  工作负荷  心理症状  同事支持/
MO NX = 1 NY = 3 BE = SD, FI GA = FU, FI PH = SY, FR PS = DI, FR TY = FI
TX = FI AL = FI KA = FI
FR GA(1, 1) GA(2, 1) GA(3, 1) BE(2, 1) BE(3, 2) BE(3, 1)
CO PAR(1) = GA(1, 1) * BE(2, 1) * BE(3, 2)
```

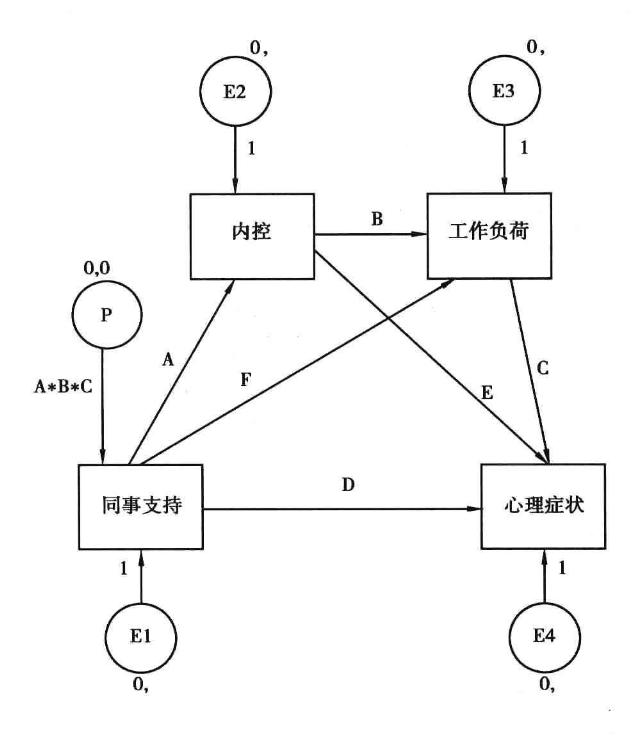


图 9-24　LISREL 双中介变量的路径设计:关联模型

```
PD
OU ME = ML IT = 250 ND = 4 BE = Indirect2. bes GA = Indirect2. gas
```

上述 LISREL 程序中,CO PAR(1)是用来建立一新的待检验参数:GA(1, 1) * BE(2, 1) * BE(3, 2),即双中介变量中介效果。

※LISREL 输出结果

TI To create a phantom var for testing indirect effect:2 mediators

Covariance Matrix

	内控	工作负荷	心理症状	同事支持
内控	0.2710			
工作负荷	0.0610	0.3960		
心理症状	0.0670	0.1160	0.2700	
同事支持	-0.0610	-0.0290	-0.0360	0.3010

Means

内控	工作负荷	心理症状	同事支持
2.3833	3.0011	2.0349	3.6382

Parameter Specifications

BETA

	内控	工作负荷	心理症状
内控	0	0	0
工作负荷	1	0	0
心理症状	2	3	0

GAMMA

	同事支持
内控	4
工作负荷	5

```
心理症状      6
        PHI
              同事支持
                 7
        PSI
              内控        工作负荷        心理症状
                8            9              10
        ALPHA
              内控        工作负荷        心理症状
               11           12              13
LISREL Estimates (Maximum Likelihood)
        BETA
              内控        工作负荷        心理症状
内控          - -          - -             - -
工作负荷      0.2131       - -             - -
             (0.0686)
              3.1048
心理症状      0.1751       0.2616          - -
             (0.0536)     (0.0435)
              3.2651       6.0166
        GAMMA
              同事支持
内控          -0.2027
             (0.0524)
              -3.8679
工作负荷      -0.0532
             (0.0651)
              -0.8161
心理症状      -0.0589
             (0.0502)
              -1.1744
        Covariance Matrix of Y and X
              内控        工作负荷        心理症状        同事支持
内控          0.2710
工作负荷      0.0610       0.3960
心理症状      0.0670       0.1160          0.2700
同事支持     -0.0610      -0.0290         -0.0360          0.3010
        Mean Vector of Eta-Variables
              内控        工作负荷        心理症状
              2.3833       3.0011          2.0349
        PHI
              同事支持
                0.3010
              (0.0241)
               12.5100
        PSI
注:This matrix is diagonal.
              内控        工作负荷        心理症状
              0.2586       0.3815          0.2258
             (0.0207)     (0.0305)        (0.0180)
              12.5100      12.5100         12.5100
```

```
Squared Multiple Correlations for Structural Equations
                内控        工作负荷      心理症状
               0.0456      0.0367       0.1637
Squared Multiple Correlations for Reduced Form
                内控        工作负荷      心理症状
               0.0456      0.0071       0.0159
        Reduced Form
               同事支持
内控           -0.2027
               (0.0524)
               -3.8679
工作负荷       -0.0963
               (0.0646)
               -1.4913
心理症状       -0.1196
               (0.0531)
               -2.2522
        ALPHA
                内控        工作负荷      心理症状
              3.1206       2.6865       1.0467
            (0.1928)     (0.3173)     (0.2707)
             16.1875       8.4660       3.8674
        ADDITIONAL PARAMETERS
               PA(1)
               -0.0113
               (0.0050)
               -2.2462
BE was written to file F:\lisrel852\Indirect2.bes
GA was written to file F:\lisrel852\Indirect2.gas
```

LISREL 报表中 Beta 与 Gamma 矩阵中的数据显示，中介效果的参数为 -0.0113 (= -0.2027 * 0.2131 * 0.2616)，其 t 检验值为 -2.2462(= -0.0113/0.005)，已达统计上的0.05 显著水平，亦即“内控”与“工作负荷”具有显著的中介效果，研究者亦可使用该参数的标准误 0.005，建立 0.05 置信区间：-0.0113 ±1.96 * 0.005。

3. 双中介变量(中介变量间无因果关系)

这是单一中介变量分析的延伸，LISREL 程序的撰写大致相同，图 9-25 中的内控与工作负荷为中介变量，但这两个中介变量间未具有因果关系。因此，本模型具有两个交互作用(A * B，C * D)，研究者可以估计两个交互作用的总效果(A * B + C * D)与两个交互作用的差异(A * B - C * D)效果。因此，本例需要两个幽灵变量，以进行这两个交互作用项的统计检验。

仍以吴丽华(2007)的研究数据为例，LISREL 的程序语法与输出报表，逐一条列于后：

※LISREL 程序设计

```
The following lines were read from file F:\lisrel852\Indirect3.LS8:

TI TO create a phantom var for testing indirect effect:2 UN-RELATED Mediators
DA NI = 4 NO = 315 NG = 1 MA = CM AP = 2
! Two additional free parameters created
```

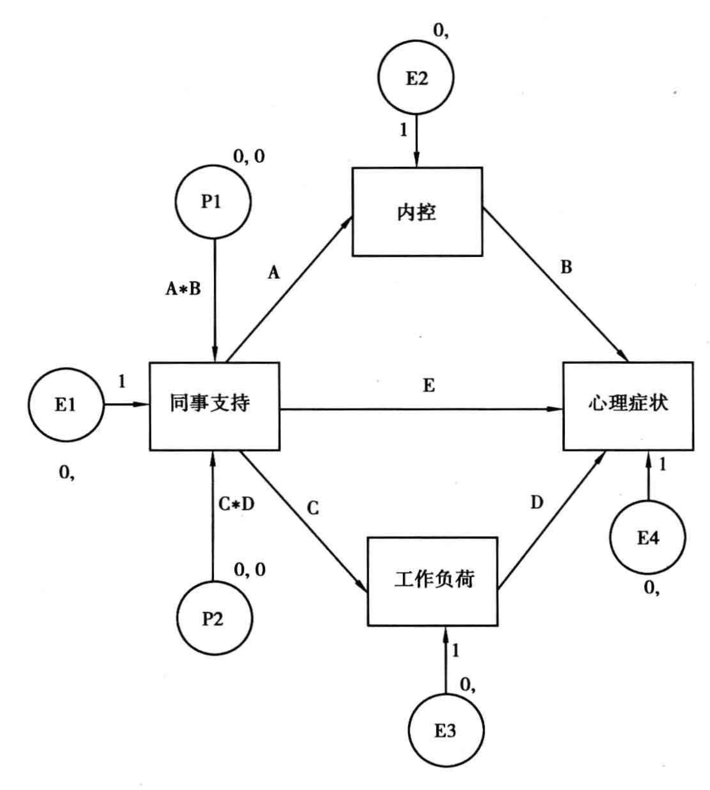


图 9-25　LISREL 双中介变量的路径设计:无关联模型

```
LA
内控　同事支持　工作负荷　心理症状
CM
.271
-.061  .301
.061  -.029  .396
.067  -.036  .116  .270
ME
2.3833  3.6382  3.0011  2.0349
SE
内控　工作负荷　心理症状　同事支持/
MO NX=1 NY=3 BE=SD, FI GA=FU, FI PH=SY, FR PS=DI, FR TY=FI
TX=FI AL=FI KA=FI
FR GA(1, 1) GA(2, 1) GA(3, 1) BE(3, 2) BE(3, 1)
! Total specific effect
CO PAR(1) =GA(1, 1) * BE(3, 1) +GA(2, 1) * BE(3, 2)
! Difference between 2 specific effect
CO PAR(2) =GA(1, 1) * BE(3, 1) -GA(2, 1) * BE(3, 2)
PD
OU ME=ML IT=250 ND=4 BE=Indirect2.bes GA=Indirect2.gas
```

上述 LISREL 程序中,CO PAR(1)、CO PAR(2)是用来建立两个待检验参数:GA(1, 1) * BE(3, 1) + GA(2, 1) * BE(3, 2),GA(1, 1) * BE(3, 1) - GA(2,1) * BE(3, 2),即双中介变量的中介效果。

※LISREL 输出结果

TI TO create a phantom var for testing indirect effect:2 UN-RELATED Mediators

Covariance Matrix

	内控	工作负荷	心理症状	同事支持
内控	0.2710			
工作负荷	0.0610	0.3960		
心理症状	0.0670	0.1160	0.2700	
同事支持	-0.0610	-0.0290	-0.0360	0.3010

Means

内控	工作负荷	心理症状	同事支持
2.3833	3.0011	2.0349	3.6382

TI TO create a phantom var for testing indirect effect:2 UN-RELATED Mediators

Parameter Specifications

BETA

	内控	工作负荷	心理症状
内控	0	0	0
工作负荷	0	0	0
心理症状	1	2	0

GAMMA

	同事支持
内控	3
工作负荷	4
心理症状	5

PHI

同事支持
6

PSI

内控	工作负荷	心理症状
7	8	9

ALPHA

内控	工作负荷	心理症状
10	11	12

TI TO create a phantom var for testing indirect effect: 2 UN-RELATED Mediators Number of Iterations = 0

LISREL Estimates (Maximum Likelihood)

BETA

	内控	工作负荷	心理症状
内控	- -	- -	- -
工作负荷	- -	- -	- -
心理症状	0.1751	0.2616	- -
	(0.0528)	(0.0428)	
	3.3150	6.1085	

GAMMA

	同事支持
内控	-0.2027
	(0.0524)
	-3.8679
工作负荷	-0.0963
	(0.0646)
	-1.4913
心理症状	-0.0589

```
               (0.0503)
               -1.1716
         Covariance Matrix of Y and X
               内控        工作负荷        心理症状        同事支持
内控           0.2710
工作负荷       0.0059      0.3960
心理症状       0.0526      0.1063          0.2649
同事支持      -0.0610     -0.0290         -0.0360          0.3010
         Mean Vector of Eta-Variables
               内控        工作负荷        心理症状
               2.3833      3.0011          2.0349
         PHI
             同事支持
               0.3010
             (0.0241)
              12.5100
         PSI
注:This matrix is diagonal.
               内控        工作负荷        心理症状
             0.2586        0.3932          0.2258
           (0.0207)      (0.0314)        (0.0180)
            12.5100       12.5100         12.5100
         Squared Multiple Correlations for Structural Equations
               内控        工作负荷        心理症状
               0.0456      0.0071          0.1478
         Squared Multiple Correlations for Reduced Form
               内控        工作负荷        心理症状
               0.0456      0.0071          0.0163
         Reduced Form
               同事支持
内控            -0.2027
               (0.0524)
                -3.8679
工作负荷        -0.0963
               (0.0646)
                -1.4913
心理症状        -0.1196
               (0.0526)
                -2.2739
         ALPHA
               内控          工作负荷          心理症状
             3.1206          3.3516            1.0467
           (0.1928)        (0.2377)          (0.2832)
            16.1875         14.1004            3.6958
         ADDITIONAL PARAMETERS
               PA(1)         PA(2)
             -0.0607       -0.0103
             (0.0224)      (0.0224)
             -2.7102       -0.4587
```

LISREL 报表显示,“内控”与“工作负荷”等两个中介变量的间接总效果为:PA(1) = -0.0607[= (-0.2027 * 0.1751) + (-0.0963 * 0.2616)],其 t 检验值为 -2.7102 $\left(\frac{-0.0607}{0.0224}\right)$,已达统计上的0.05 显著水平,亦即经由“内控”与“工作负荷”等两个中介变

量，具有显著的中介总效果。另外，经由内控与工作负荷中介变量的间接差异效果：PA(2) $= -0.00103[=(-0.2027*0.1751)-(-0.0963*0.2616)]$，其 t 检验值为 -0.4587 $\left(\frac{-0.0103}{0.0224}\right)$，未达统计上的0.05 显著水平，亦即经由"内控"与经由"工作负荷"的间接效果，两者间未具有显著的差异。

(三)Bayesian SEM 自订参量

Amos 虽然无法使用幽灵变量法进行中介效果值的显著性检验，但可利用其 Bayesian SEM 自订参量，进行中介效果值的显著性检验并计算置信区间。请注意利用 SEM 进行中介效果值的显著性检验，研究者必须事先根据理论或过去的研究结果，决定所提议的中介效果模型是完全中介还是部分中介。完全中介时，在路径图中就不须画出直接效果的路径，此时中介效果值是两个回归系数的乘积；部分中介时，在路径图中就须画出直接效果的路径，此时中介效果值是两个净回归系数的乘积。

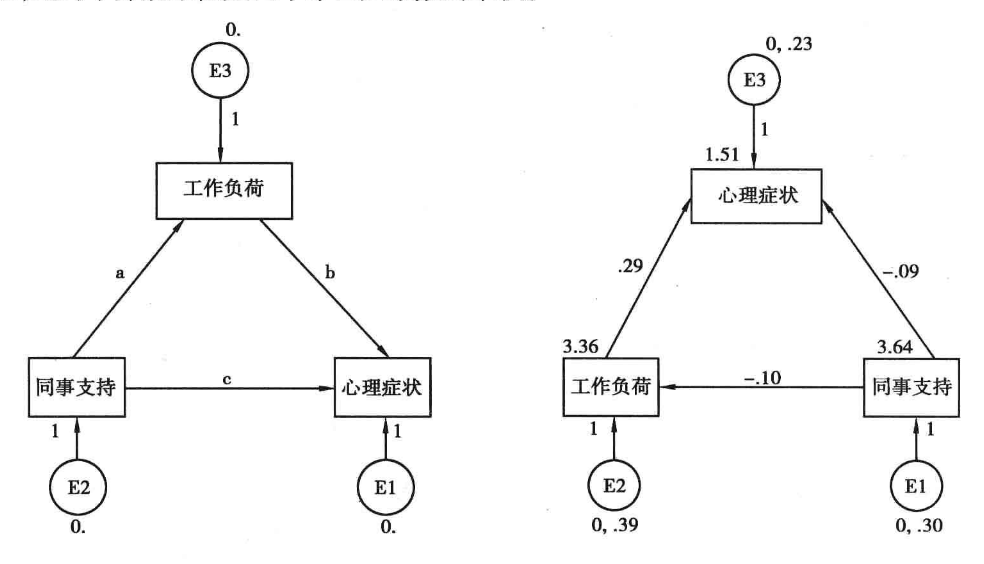


图 9-26　Amos 单一中介变量的路径图　　　图 9-27　Amos 单中介变量的路径系数

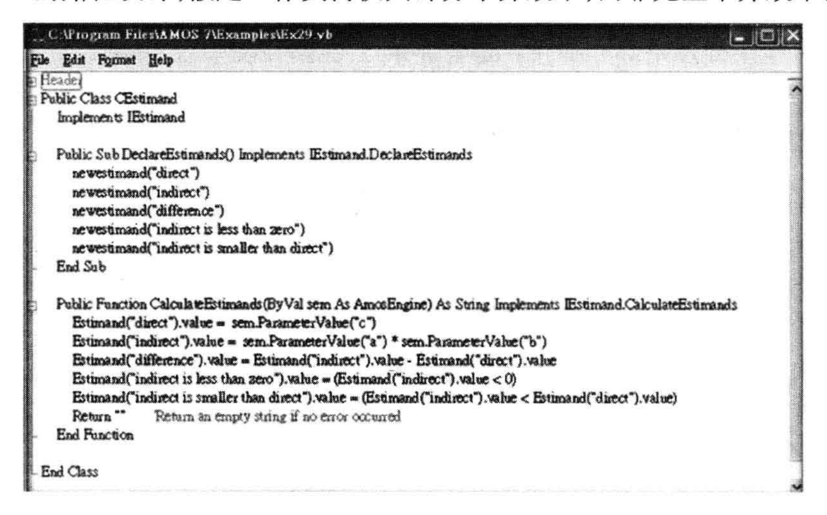
根据 James，Mulaik 与 Brett(2006)的建议：当您的提议模型为完全中介效果时，模型中($Y=\beta_{10}+\beta_{11}X+\varepsilon_1$，$Me=\beta_{20}+\beta_{21}X+\varepsilon_2$)的两个回归系数 β_{11}，β_{21} 均须达 0.05 显著水平，但这两个回归系数的交乘积值($\beta_{11}*\beta_{21}$)及预测变量(X)与效标(Y)间相关(r_{XY})的差异值($\beta_{11}*\beta_{21}-r_{XY}$)不能达 0.05 显著水平，完全中介的提议模型才能获得支持。当您的提议模型为部分中介效果时，模型中($Y=\beta_{30}+\beta_{31}X+\beta_{32}Me+\varepsilon_3$)的两个净回归系数($\beta_{31}$，$\beta_{32}$)均须达 0.05 显著水平，而且预测变量对中介变量(Me)的回归系数(β_{21})亦须达 0.05 显著水平，部分中介效果的提议模型才能获得支持。这些基本前提，乃是研究者事先须来厘清的工作。

以下，介绍如何利用 Amos Bayesian SEM 计算中介效果与相关参数的统计检验。Amos 的操作步骤，须先划出如图 9-26 的提议模型(路径图中须标注路径参数 a，b，c)、连接数据文件、使用 Amos 的[Analyze]功能表单下的[Calculate Estimates]计算出模型中参数(a，b，c)的值。

其次,使用[Analyze]功能表单下的[Bayesian Estimation]执行贝氏估计法,等 Bayesian SEM 表单上出现笑脸后,按下暂停钮暂停抽样,并按下进行自订参数的估计(Custom Estimation),等 Amos 出现自订参数估计的程序设计窗口后,在程序设计窗口中撰写中介效果估计值的相关程序,请参见图 9-28 窗口中的程序设计。为便于比较不同方法(幽灵变量法 vs Bayesian SEM 分析法),以下仍依中介变量的个数与因果关系:单一中介变量、关联双中介变量与无关联双中介变量等三类中介形式,逐一利用吴丽华(2007)的研究数据为例说明如下。

1. 单一中介变量

图 9-26 的路径设计,假定工作负荷仅具部分中介效果,而非完全中介效果。



```
C:\Program Files\AMOS 7\Examples\Ex29.vb
File Edit Format Help
Header
Public Class CEstimand
    Implements IEstimand

    Public Sub DeclareEstimands() Implements IEstimand.DeclareEstimands
        newestimand("direct")
        newestimand("indirect")
        newestimand("difference")
        newestimand("indirect is less than zero")
        newestimand("indirect is smaller than direct")
    End Sub

    Public Function CalculateEstimands(ByVal sem As AmosEngine) As String Implements IEstimand.CalculateEstimands
        Estimand("direct").value = sem.ParameterValue("c")
        Estimand("indirect").value = sem.ParameterValue("a") * sem.ParameterValue("b")
        Estimand("difference").value = Estimand("indirect").value - Estimand("direct").value
        Estimand("indirect is less than zero").value = (Estimand("indirect").value < 0)
        Estimand("indirect is smaller than direct").value = (Estimand("indirect").value < Estimand("direct").value)
        Return ""    Return an empty string if no error occurred
    End Function

End Class
```

图 9-28 Amos Bayesian SEM 中介效果的程序设计:自订参量

Custom Estimands

File Edit View Help

Numeric Estimands	Mean	S.D.	C.S.	95% Lower bound	95% Upper bound
direct	-0.095	0.051	1.001	-0.194	0.002
indirect	-0.027	0.019	1.001	-0.066	0.008
difference	0.067	0.053	1.001	-0.036	0.173

Dichotomous Estimands	P	P1	P2	P3
indirect is less than zero	0.935	0.950	0.922	0.932
indirect is smaller than direct	0.104	0.133	0.081	0.098

图 9-29 Amos Bayesian SEM 中介效果的输出报表

由图 9-29 的报表知,中介效果的平均值为 -0.027,在全部 MCMC 样本中,此值小于 0 的概率为 0.935,前 1/3(p1)、2/3(p2)与后 1/3(p3)的概率亦皆甚为接近(0.922 ~ 0.950),可见 p 值相当稳定,反映出中介效果为负的概率;而其 0.95 置信区间为 -0.066 ~

0.008,因置信区间包含0,显示出此中介效果值未达0.05 显著水平。此项分析结果与前述的幽灵变量法的结果类似。至于直接效果与间接效果间的差异检验,因0.95 置信区间包含0(-0.036~0.173),亦未达0.05 显著水平。

2. 双中介变量(中介变量有关联)

这是单一中介变量分析的延伸,Amos 贝氏 SEM 程序的撰写大致相同,图9-30 中的内控与工作负荷均为中介变量,且这两个中介变量间具有因果关系,图中相关的路径系数分别标注为A,B,C,D,E,F。

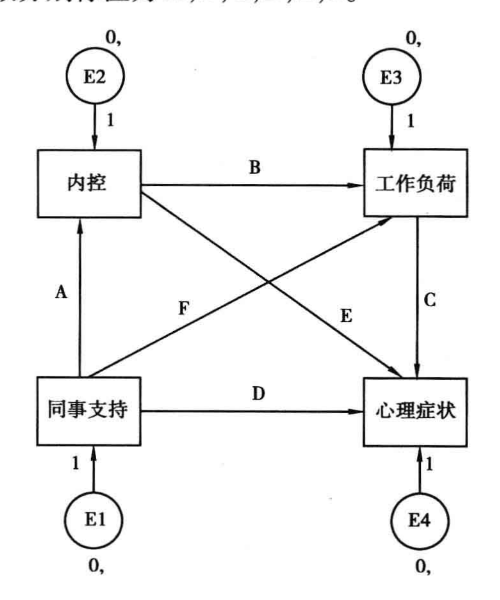


图9-30 Amos 双中介变量的路径设计图:中介变量间具有因果关系

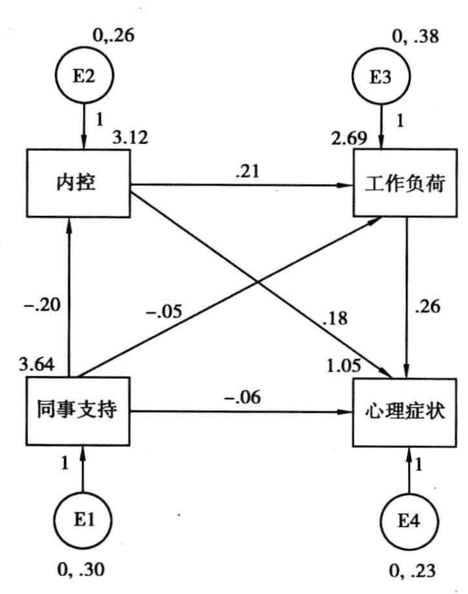


图9-31 Amos 双中介变量的路径系数

研究者使用 Amos 的[Analyze]功能表单下的[Calculate Estimates]计算出模型中参数的值;图9-31 是中双中介变量的路径系数,是 Amos 估计出来的参数值。之后,使用[Analyze]功能表单下的[Bayesian Estimation]执行贝氏估计法,等 Bayesian SEM 表单上出现笑脸后,按下暂停钮暂停抽样,并按下进行自订参数的估计,等 Amos 出现自订参数估计的程序设计窗口后,在程序设计窗口中撰写中介效果估计值的相关程序。Amos 的相关 Bayesian SEM 程序设计列述于后,以供研究者参考。

```
#Region "Header"
Imports System
Imports Microsoft. VisualBasic
Imports AmosEngineLib
Imports AmosEngineLib. AmosEngine. TMatrixID
Imports AmosExtensions. CustomEstimand
Imports PBayes
#End Region
Public Class CEstimand
    Implements IEstimand
    Public Sub DeclareEstimands( ) Implements IEstimand. DeclareEstimands
      newestimand ("直接效果")
```

```
        newestimand ("间接效果")
        newestimand ("difference")
        newestimand ("间接效果 is less than zero")
        newestimand ("间接效果 is smaller than direct")

    End Sub
    Public Function CalculateEstimands (ByVal sem As AmosEngine) As
String Implements IEstimand. CalculateEstimands
    Estimand ("直接效果"). value = sem. ParameterValue ("d")
    Estimand ("间接效果"). value = sem. ParameterValue ("a") * sem.
ParameterValue ("b") * sem. ParameterValue ("c")
    Estimand ("difference"). value = Estimand ("间接效果"). value -
Estimand ("直接效果"). value
    Estimand ("间接效果 is less than zero"). value = (Estimand ("间接效果").
value < 0)
    Estimand ("间接效果 is smaller than direct"). value = (Estimand ("间接
效果"). value < Estimand ("直接效果"). value)
    Return ""      'Return an empty string if no error occurred
    End Function
End Class
```

图 9-32 窗口内的结果，是 Amos Bayesian SEM 双中介效果的输出报表。

由图 9-32 的报表知，双中介效果的平均值为 -0.0113，就全部 MCMC 样本中，此值小于0 的概率为 0.9983，前 1/3(p1)、2/3(p2) 与后 1/3(p3) 的概率亦皆接近于 1(0.9978 ~0.9988)，可见 p 值相当稳定，反映出中介效果肯定为负的；而其 0.95 置信区间为 -0.00231 ~0.0031，因不包含 0，显示出双中介效果值达 0.05 显著水平。此项分析结果与前述的幽灵变量法的结果亦非常接近。至于直接效果与间接效果间的差异检验，因0.95 置信区间包含 0(-0.052 ~0.145)，亦未达 0. 05 显著水平；而由其概率等于 0.174 来看，直接效果大于间接效果的概率满高的。

3. 双中介变量(中介变量无关联)

这是单一中介变量分析的延伸，Amos[Bayesian SEM]程序的撰写大致相同，图 9-33 中的内控与工作负荷为中介变量，但这两个中介变量间未具有因果关系。因此，本模型具有两个交互作用(A * B，C * D)，研究者可以获得一个交互作用的总效果与两个交互作用的差异效果。

研究者须先使用 Amos 中[Analyze]功能表单下的[Calculate Estimates]计算出模型中参数的值。图 9-34 中双中介变量的路径系数，是 Amos 估计出来的参数值。之后，使用[Analyze]功能表单下的[Bayesian Estimation]执行贝氏估计法，等 Bayesian SEM 表单上出现笑脸后，按下暂停钮暂停抽样，并按下进行自订参数的估计，等 Amos 出现自订参数估计的程序设计窗口后，在程序设计窗口中撰写中介效果估计值的相关程序。

兹将 Amos 的相关 Bayesian SEM 程序设计列述于下，以供研究者参考。

```
#Region "Header"
Imports Systeam
Imports Microsoft. VisualBasic
Imports AmosEngineLib
Imports AmosEngineLib. AmosEngine. TMatrixID
```

Custom Estimands

Numeric Estimands	Mean	S.E.	S.D.	C.S.	Skewness	Kurtosis	Min
直接效果	-0.060	0.002	0.051	1.001	-0.056	0.227	-0.284
间接效果	-0.011	0.000	0.005	1.000	-0.685	0.592	-0.035
difference	0.049	0.002	0.051	1.001	0.063	0.203	-0.150

Dichotomous Estimands	P	P1	P2	P3
间接效果　ess than zero	0.998	0.999	0.998	0.997
间接效果 : smaller than direct	0.167	0.188	0.149	0.163

图 9-32　Amos Bayesian SEM 双中介效果的输出报表:双中介有关联模型

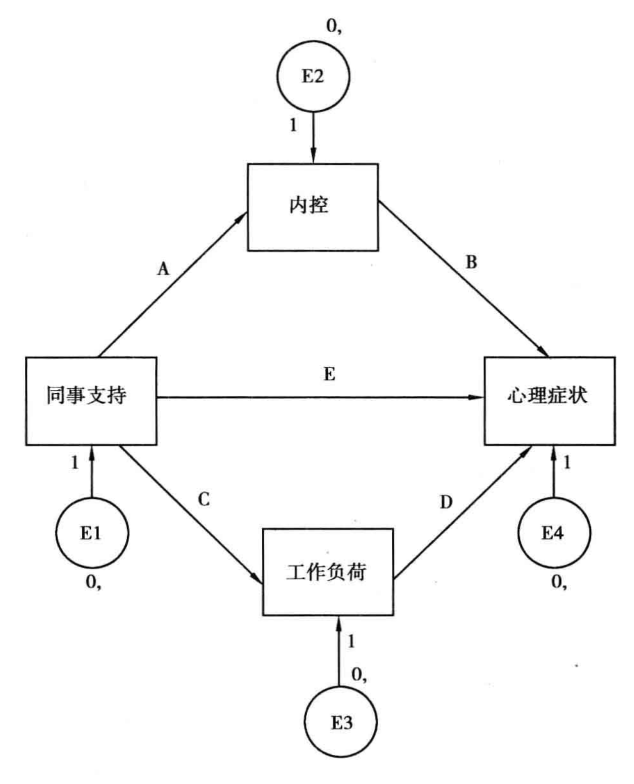


图 9-33　Amos 双中介变量的路径设计图

```
Imports AmosExtensions. CustomEstimand
Imports PBayes
#End Region
Public Class CEstimand
    Implements IEstimand
    Public Sub DeclareEstimands( ) Implements IEstimand. DeclareEstimands
```

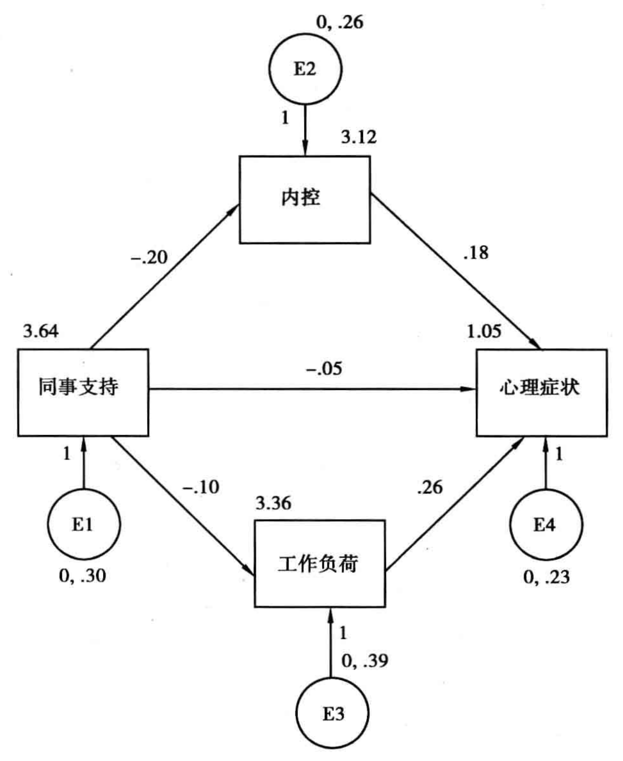


图 9-34　Amos 双中介变量的路径系数

```
        newestimand("direct")
        newestimand("indirect1")
        newestimand("indirect2")
        newestimand("Total Effect")
        newestimand("difference")
        newestimand("indirect1 is less than zero")
        newestimand("indirect2 is less than zero")
        newestimand("Total Effect is less than zero")
        newestimand("Difference is less than zero")
        newestimand("Total indirect is smaller than direct")
    End Sub
    Public Function CalculateEstimands(ByVal sem As AmosEngine) As
String Implements IEstimand. CalculateEstimands
        Estimand("direct").value = sem.ParameterValue("E")
        Estimand("indirect1").value = sem.ParameterValue("a") * sem.
ParameterValue("b")
        Estimand("indirect2").value = sem.ParameterValue("C") * sem.
ParameterValue("D")
        Estimand("Total Effect").value = Estimand("indirect1").value +
Estimand("indirect2").value
        Estimand("Difference").value = Estimand("indirect1").value -
Estimand("indirect2").value
        Estimand("indirect1 is less than zero").value = (Estimand("indirect1").
value < 0)
```

```
    Estimand("indirect2 is less than zero").value = (Estimand("indirect2").
value < 0)
        Estimand("Total Effect is less than zero").value = (Estimand("Total
Effect").value < 0)
        Estimand("Difference is less than zero " ) . value =
(Estimand("Difference").value < 0)
        Estimand("Total indirect is smaller than direct").value =
(Estimand("Total Effect").value < Estimand("direct").value)
        Return ""      'Return an empty string if no error occurred
    End Function
End Class
```

图9-35 窗口内的结果,是 Amos Bayesian SEM 双中介效果的输出报表。由图9-35 报表知,在无关联模型下,双中介总效果的平均值为 -0.062,就全部 MCMC 样本来看,此值小于0 的概率为0.997;而其0.95 置信区间为 -0.107 ~ -0.019,因置信区间不包含0,显示出双中介总效果值达0.05 显著水平。至于双中介差异效果的平均值为 -0.010[-0.036-(-0.026)],此值小于0 的概率为0.677,其0.95 置信区间为 -0.057 ~ 0.037,因置信区间包含0,显示出双中介差异效果值未达0.05 显著水平。此项分析结果与前述的幽灵变量法的结果亦非常接近。就全部 MCMC 样本来看,双中介总效果值小于直接效果值的概率为0.522。

Custom Estimands

File Edit View Help

Numeric Estimands	Mean	S.D.	C.S.	95% Lower bound	95% Upper bound
direct	-0.058	0.051	1.001	-0.159	0.039
indirect1	-0.036	0.014	1.001	-0.067	-0.011
indirect2	-0.026	0.018	1.001	-0.063	0.008
Total Effect	-0.062	0.022	1.001	-0.107	-0.019
difference	-0.010	0.024	1.001	-0.057	0.037

Dichotomous Estimands	P	P1	P2	P3
indirect1 is less than zero	0.999	0.999	0.999	1.000
indirect2 is less than zero	0.932	0.929	0.926	0.942
Total Effect is less than zero	0.997	0.998	0.997	0.996
Difference is less than zero	0.677	0.669	0.681	0.680
Total indirect is smaller than direct	0.522	0.525	0.531	0.510

图9-35　Amos Bayesian SEM 双中介效果的输出报表:无关联模型

四、Bootstrap 法

Bootstrapping 统计分析不受抽样分布形态的限制,在电脑 CPU 速率日益强大的今日,Bootstrapping 的统计分析受到不少应用者的青睐。间接效果值(a * b)的 Bootstrapping 方法道理很简单,其 bootstrap 抽样分布及 a * b 的估计值,可经由下列四大步骤求得:

1. 使用原始数据作为抽样的总体，用“放回抽样法”从总体中随机抽取 N 个 bootstrap 样本点，建立一个 bootstrap 样本；
2. 计算这个 bootstrap 样本的 a * b 值，并将之存文件；
3. 重复步骤 1 ~ 2 若干次（如 1 000 次）；
4. 利用这些 bootstrap 所得的 A * b 值，建立 a * b 的抽样分布，并计算（α/2）* 100% 与（1 - α/2）* 100% 的百分位数与该 bootstrap 抽样分布的平均数、标准差。

研究者如欲使用 Amos 的 Bootstrap 法计算直接效果、间接效果与总效果，请参照以下步骤进行，堪称便捷。首先，绘制所需的路径图与数据文件连接。接着点选分析属性窗口中的[Output]，在打开的窗口中勾选[Indirect, direct & total effects]（参见图 9-36）。

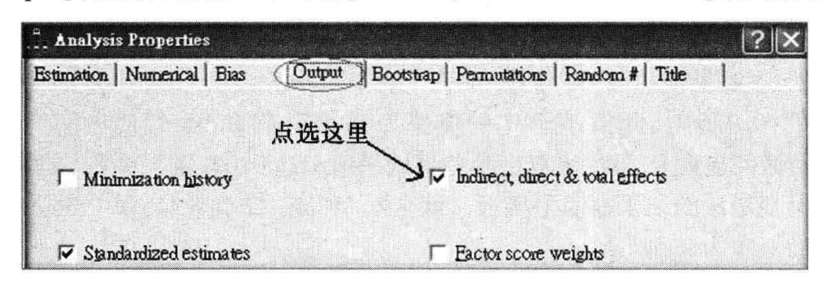


图 9-36　Amos 分析属性 Output 的设定

其次，再点选分析属性窗口中的[Bootstrap]，在打开的窗口中勾选[Perform bootstrap]、[Percentile confidence intervals]、[Bias-corrected confidence intervals]与[Bootstrap Ml]（参见图 9-37）。最后点选[Calculate Estimates]，即可获得相关的统计量。

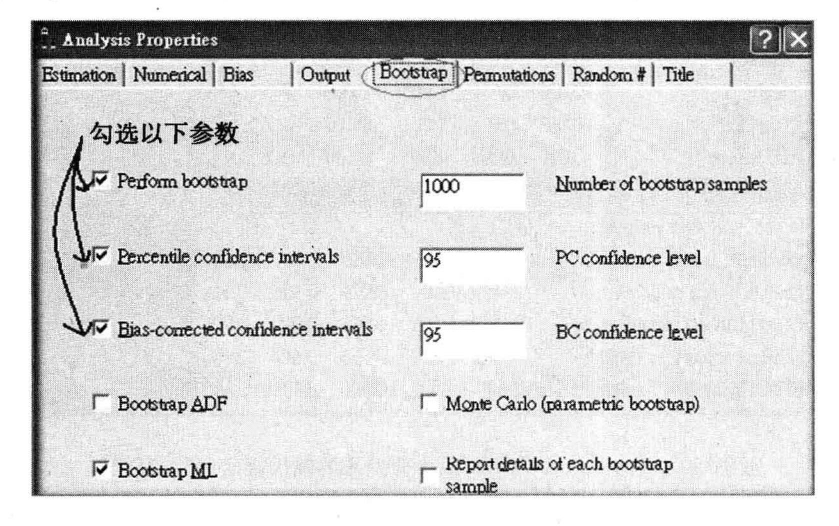


图 9-37　Amos 分析属性 Bootstrap 的设定

本法的 Amos 应用实例解说，请看本章第六节的说明，不在此赘述。

第五节　条件式中介效果

一、定义与类型

条件式中介效果也称为条件式间接效果(conditional indirect effect),或称为调节型中介效果(moderated mediation),指的是间接效果的大小与方向根据调节变量的某一层次或某特定数值而定(Muller, Judd, & Yzerbyt, 2005; Preacher, Rucker, & Hayes, 2007),亦即中介或间接效果的大小受到调节变量的控制。因此,条件式间接或中介效果的研究,旨在探讨间接或中介效果如何发生与什么时候会发生。Preacher, Rucker 与 Hayes(2007)整理出五个条件式中介效果的模型,最简单的模型是自变量就是调节变量,该模型中,X 是起始变量也是调节变量。当 X 与 M 交互作用的效果接近于 0,b_2 亦会接近于 0 时,此时条件式间接效果会简化为 a_1b_1,对任何调节变量的层次或个案均为 a_1b_1。

(一)模型一:条件式间接效果的路径模型

如图 9-38 所示:

此模型的相关回归方程为:

$$M = a_0 + a_1X + e$$

$$Y = b_0 + c'X + (b_1 + b_2X)M + e$$

其条件式间接效果的点估计与二阶变异误为:点估计为 $a_1(b_1 + b_2X)$,而其二阶变异误为

$$(\hat{b}_1 + \hat{b}_2X)^2 s_{\hat{a}_1}^2 + (\hat{a}_1^2 + s_{\hat{a}_1}^2)(s_{\hat{b}_1}^2 + 2s_{\hat{b}_1\cdot\hat{b}_2}X + s_{\hat{b}_2}^2X^2)$$

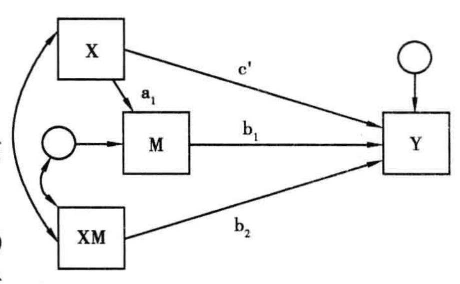


图 9-38　条件式中介效果:模型一

利用上述间接效果的点估计与二阶变异量,就可计算 t 值与 p 值。为便于研究者计算条件式间接效果,其余的模型二 ~ 五的条件式间接效果的理论模型、点估计与二阶变异误(将之开根号,即为标准误),亦摘述如下,以便研究者计算临界比(Critical ratio)。

(二)模型二:路径设计

如图 9-39 所示,路径 a 受到 W 变量的调节:点估计为 $(\hat{a}_1 + \hat{a}_3W)\hat{b}_1$,而其二阶变异误为:

$$(\hat{a}_1 + \hat{a}_3W)^2 s_{\hat{b}_1}^2 + (\hat{b}_1^2 + s_{\hat{b}_1}^2)(s_{\hat{a}_1}^2 + 2s_{\hat{a}_1\cdot\hat{a}_3}W + s_{\hat{a}_3}^2W^2)$$

(三)模型三:路径设计

如图 9-40 所示,路径 b 受到 W 变量的调节:点估计为 $\hat{a}_1(\hat{b}_1 + \hat{b}_3W)$,而其二阶变异误为:

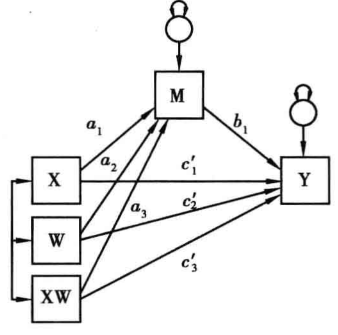
$$(\hat{b}_1 + \hat{b}_3W)^2 s_{\hat{a}_1}^2 + (\hat{a}_1^2 + s_{\hat{a}_1}^2)(s_{\hat{b}_1}^2 + 2s_{\hat{b}_1\cdot\hat{b}_3}W + s_{\hat{b}_3}^2W^2)$$

(四)模型四:路径设计

如图 9-41 所示,路径 a 受到 W 变量的调节,路径 b 受到 Z 变量的调节:点估计为

$(\hat{a}_1+\hat{a}_3W)(\hat{b}_1+\hat{b}_3Z)$，而其二阶变异误为：

$$(\hat{a}_1+\hat{a}_3W)^2(s^2_{\hat{b}_1}+2s_{\hat{b}_1\cdot\hat{b}_3}Z+s^2_{\hat{b}_1}Z^2)+(\hat{b}_1+\hat{b}_3Z)^2(s^2_{\hat{a}_1}+2s_{\hat{a}_1\cdot\hat{a}_3}W+s^2_{\hat{a}_3}W^2)$$
$$+(s^2_{\hat{b}_1}+2s_{\hat{b}_1\cdot\hat{b}_3}Z+s^2_{\hat{b}_3}Z^2)(s^2_{\hat{a}_1}+2s_{\hat{a}_1\cdot\hat{a}_3}W+s^2_{\hat{a}_3}W^2)$$



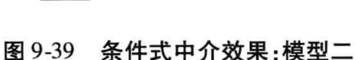


图 9-39　条件式中介效果：模型二

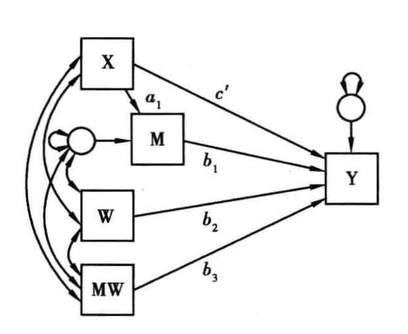


图 9-40　条件式中介效果：模型三

（五）模型五：路径设计

如图 9-42 所示，路径 a & b 均受到 W 变量的调节。点估计为：$(\hat{a}_1+\hat{a}_3W)(\hat{b}_1+\hat{b}_2Z)$，而其二阶变异误为：

$$(\hat{b}_1+\hat{b}_2W)^2(s^2_{\hat{a}_1}+2s_{\hat{a}_1\cdot\hat{a}_3}W+s^2_{\hat{a}_3}W^2)+(\hat{a}_1+\hat{a}_3W)^2(s^2_{\hat{b}_1}+2s_{\hat{b}_1\cdot\hat{b}_2}W+s^2_{\hat{b}_2}W^2)$$
$$(s^2_{\hat{a}_1}+2s_{\hat{a}_1\cdot\hat{a}_3}W+s^2_{\hat{a}_3}W^2)(s^2_{\hat{b}_1}+2s_{\hat{b}_1\cdot\hat{b}_2}Z+s^2_{\hat{b}_2}W^2)$$

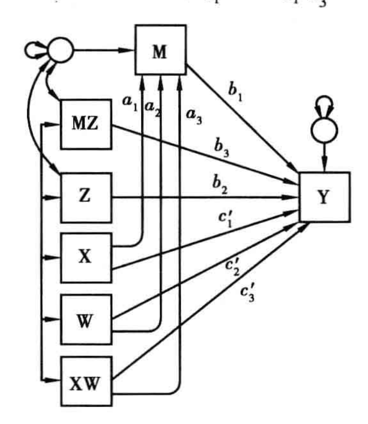


图 9-41　条件式中介效果：模型四

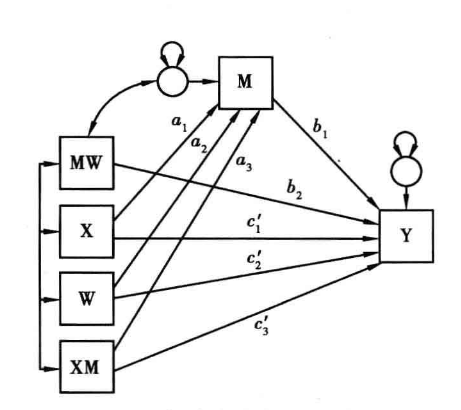


图 9-42　条件式中介效果：模型五

另外，依照 Muller，Judd 与 Yzerbyt（2005）的主张，欲检验调节型（或称条件化）中介效果（moderated mediation），须采用以下三个回归方程：

$$Y=\beta_{40}+\beta_{41}X+\beta_{42}Mo+\beta_{43}XMo+\varepsilon_4$$
$$Me=\beta_{50}+\beta_{51}X+\beta_{52}+Mo+\beta_{53}XMo+\varepsilon_3$$
$$Y=\beta_{60}+\beta_{61}X+\beta_{62}Mo+\beta_{63}XMo+\beta_{64}Me+\beta_{65}MeMo+\varepsilon_6$$

式中 Mo 代表调节变量，Me 代表中介变量。此种调节型中介效果，意味着 X 与 Y 间的间接效果的大小须视调节变量而定。欲证明调节型中介效果的存在，首先要检验 H_01：$\beta_{41}=0$；H_02：$\beta_{43}=0$（请参见第一个回归方程序），期望能推翻 H_01（β_{41} 达显著水平）；但希

望能接纳 $H_0 2$（显示 X 与 Mo 间无交互作用），β_{43}代表整体处理效果的交互作用。其次，在第二、三个回归方程中，则期望回归系数 β_{53}与 β_{64}或 β_{51}与 β_{65}达到显著水平，而符合这项期望的后果就是 β_{63}（代表净处理效果的交互作用）会达显著水平。

Muller, Judd 与 Yzerbyt（2005）认为除了调节型中介效果的外，尚有中介型调节作用（mediated moderation）。欲证明中介型调节作用的存在，首先要检验 H_0: $\beta_{43} = 0$（请参见第一个回归方程），期望能推翻 H_0（β_{43}达显著水平），希望 X 与 Mo 间具有整体交互作用。其次，在第二、三个回归方程中，则期望回归系数 β_{53}与 β_{64}或 β_{51}与 β_{65}达显著水平，而显示出调节作用的强度会因模型中的中介变量而降低。下面以图 9-43 的路径系数说明两者的分野所在。

调节型中介效果与中介型调节作用主要由 β_{43}绝对值与 β_{63}绝对值的差异（$|\beta_{43} - \beta_{63}|$）而定（参见图 9-43 中相对应的参数），当其绝对值为 0 时，调节型中介效果与中介型调节作用均不会发生，亦即仅有直接效果存在。当其绝对值大于 0 时，即为中介型调节作用，此时研究者所关心的是何时会产生交互作用，处理效果的交互作用是否会因中介变量的介入而下降。当其绝对值小于 0 且 β_{43}的效果等于 0 时，即为调节型中介效果，此时研究者所关心的是何时会产生中介效果，亦即处理变量的中介效果是否会因调节变量而有不同。请注意，假如 β_{43}效果等于 0 的基本假设放宽时，Muller, Judd 与 Yzerbyt（2005）认为$|\beta_{43}| - |\beta_{63}|$的差异值大于 0，即可称为调节型中介效果。在此较宽松的定义下，中介型调节作用皆是调节型中介效果，但反之则不然。



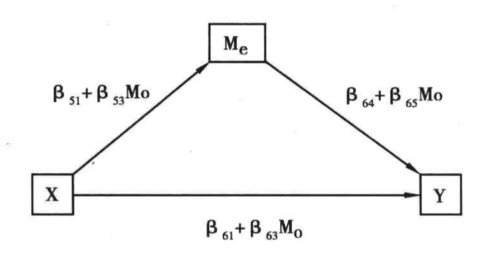


图 9-43　调节型中介效果与中介型调节作用

二、条件式中介效果实例解说：以第五模型为例

本应用实例取自 Petty 等人（1993）对于正负向情绪（MOOD）、正向思考（POS）、认知需求（NFC）与态度改变（ATT）关系的研究（Muller, Judd & Yzerbyt, 2005）。他们认为正负向情绪经由中介变量正向思考对态度改变的中介效果，会因调节变量认知需求的高低而有交互作用。因此，MOOD 被视为实验处理的自变量，ATT 被视为实验处理的依变量，POS 被视为中介变量，NFC 被视为调节变量。他们的数据可自以下网址取得 http://www.psp.ucl.ac.be/mediation/。

图 9-44 的 SPSS 数据文件中，MOODNFC，POSNFC 是交互作用乘积项。

modmed.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : OBS　1

	OBS	MOOD	NFC	MOODNFC	POS	POSNFC	ATT
1	1	-1	3.07579	-3.07579	6.25	19.2238	18.5455
2	2	-1	2.59489	-2.59489	-5.07	-13.1485	-16.9684
3	3	-1	-1.00952	1.00952	.93	-.9435	-5.9578
4	4	-1	-.43824	.43824	2.87	-1.2563	-7.2256
5	5	-1	.21788	-.21788	-16.56	-3.6081	-26.7000
6	6	-1	.43842	-.43842	-13.14	-5.7593	-24.1241
7	7	-1	-1.09646	1.09646	-.95	1.0451	-28.0378
8	8	-1	-.57737	.57737	-13.77	7.9523	-9.4845
9	9	-1	1.75978	-1.75978	-5.82	-10.2359	10.8746
10	10	-1	-.42134	.42134	.33	-.1381	18.2762

图 9-44　SPSS 数据文件

数据来源：Petty 等人（1993）研究数据（N = 100）

研究者欲分析调节型中介效果，主要的方法有三：SPSS 线性回归分析，Bootstrap 法，SEM 分析法。以下依此分类逐一说明。

（一）SPSS 线性回归分析法

Petty 等人（1993）的研究中，认为正负向情绪（MOOD）经由中介变量正向思考（POS）对态度改变（ATT）的中介效果，会因调节变量认知（NFC）需求的高低而有交互作用。因此，MOOD 被视为实验处理的自变量（X），ATT 被视为实验处理的依变量（Y），POS 被视为中介变量（Me），NFC 被视为调节变量（Mo）。有了这些基本概念后，就容易说明如何利用 SPSS 的回归分析，进行前述 Muller，Judd 与 Yzerbyt（2005）的调节型中介效果分析。在建立了图 9-44 的 SPSS 数据文件后，就可打开如图 9-45 SPSS 线性回归分析输入窗口，进行相关变量的键入，兹将分析的步骤简介如下。

第一，在图 9-45 的[Independent(s)]窗口中，放进 MOOD，NFC，MOODNFC（交互作用项）等三个预测变量，效标[Dependent]为 ATT，以检验前述第一个回归方程式。注意此阶段，未将与中介变量相关的变量（POS & POSNFC）放入回归方程式中。检验结果如表 9-11 所示。

表 9-11　SPSS 回归分析报表：第一个回归方程式

Modal	Unstanardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta.	t	Sig.
1　(Constant)	1.959	1.543		1.269	.207
MOOD	6.814	1.543	.408	4.415	.000
NFC	1.268	1.135	.106	1.117	.267
MOODNFC	-.691	1.135	-.058	-.609	.544

a. Dependent Variable：ATT.

表 9-11 结果显示，MOOD 对于依变量（ATT）的直接处理效果达 0.05 显著水平

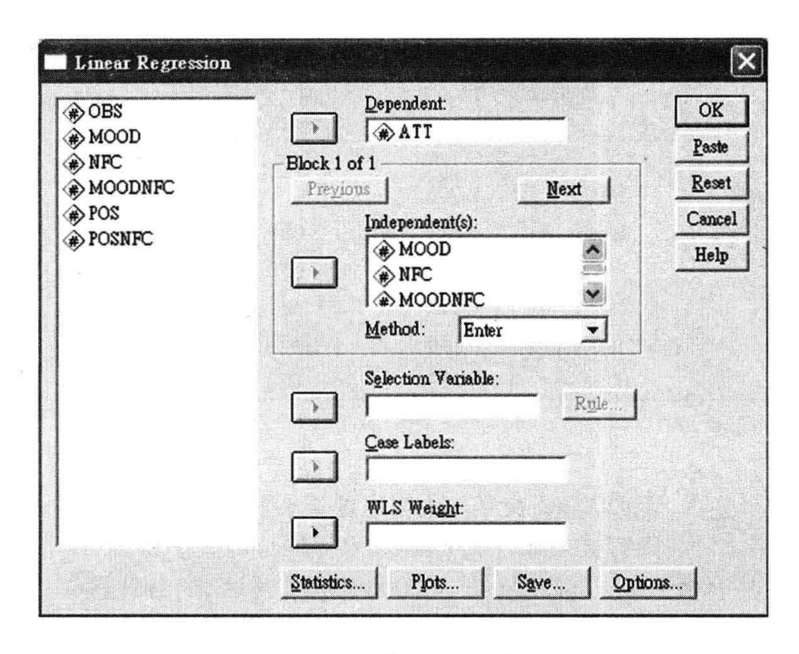


图 9-45　SPSS 线性回归分析的输入窗口

(β_{41} = 6.814, p = 0.000)，而此直接效果与认知需求变量 NFC 没有交互作用(β_{43} = −0.691, p = 0.544)。因此，这是典型调节型中介效果(prototypic moderated mediation)的基本条件。有些研究者认为先有直接效果，才能谈间接效果。其实，当直接效果为 0 时，表示完全中介。因此，也有些研究者认为此条件并不一定需要成立。

第二，放进 MOOD, NFC, MOODNFC(交互作用项)等三个预测变量，效标为 POS，以检验前述 Muller, Judd 与 Yzerbyt(2005)的第二个回归方程式。结果如表 9-12 所示。

表 9-12　SPSS 回归分析报表：第二个回归方程式

Modal	Unstanardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta.	t	Sig.
1 (Constant)	.040	.697		.058	.954
MOOD	4.336	.697	.524	6.219	.000
NFC	.767	.513	.129	1.496	.138
MOODNFC	1.257	.513	.212	2.450	.016

a. Dependent Variable: POS.

表 9-12 结果显示，MOOD 对于 POS 具有显著的处理效果(β_{51} = 4.336, p = 0.000)，MOODNFC 交互作用项达显著水平(β_{53} = 1.257, p = 0.016)，符合调节型中介效果的基本要件，MOOD 中介效果的大小为中介变量 NFC 的函数。

第三，放进 MOOD, NFC, MOODNFC, POS, POSNFC 等五个预测变量，效标为 ATT，以检验前述第三个回归方程式。结果如表 9-13 所示。

表 9-13 SPSS 回归分析报表：第三个回归方程式

Modal	Unstanardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta.	t	Sig.
1 (Constant)	1.939	1.293		1.500	.137
MOOD	1.480	1.547	.089	.957	.341
NFC	.356	.972	.030	.366	.715
MOODNFC	−2.169	1.027	−.181	−2.112	.037
POS	1.248	.189	.618	6.613	.000
POSNFC	−0.36	.129	−.023	−.278	.781

a. Dependent Variable：ATT.

表 9-13 结果显示，MOODNFC 交互作用项达显著水平（$\beta_{63} = -2.169, p = 0.037$），亦即排除间接效果之后其净直接效果（residual direct effect）的大小得视 NFC 的高低而定。此外，POS 亦达显著水平（$\beta_{64} = 1.248, p = 0.000$），与前述 β_{53} 的显著性结果（$p = 0.016$）刚好符合调节型中介效果的所有基本要件。因为 $|-0.691| - |-2.169| = -1.478$，而且 $\beta_{43} = -0.691$（$p = 0.544$），反映出经由 POS 的间接效果大小得视调节变量 NFC 的高低而定，这是典型的调节型中介效果的例子：研究者所关心的是处理变量的中介效果，是否会因调节变量而有所不同。

最后，研究者可以利用前述的 Freedman 与 Schatzkin（1992）公式，利用净直接效果 1.478[−0.691 −（ −2.169）] 进行整体调节型中介效果的统计检验：

$$t = 1.478 / \sqrt{1.135^2 + 1.027^2 - 2 * 1.135 * 1.027\sqrt{1 - .542^2}}$$
$$= 1.478/.3574 = 4.136 (p < .05)$$

因此，从净直接效果来看，可以间接看出整体调节型中介效果达到了统计上的 0.05 显著水平。

（二）Bootstrap 法

传统的中介效果分析，会受到样本大小及正态分布的限制，许多研究者都推荐使用 Bootstrap 法（Preacher, Rucker & Hayes, 2007）。为便利研究者使用 Bootstrap 法进行调节型中介效果分析，他们提供以下 SPSS Macro 程序，非常简便。本程序在 SPSS 未关闭以前，只能执行一次。执行本程序后，即可运用它来分析调节型中介效果。此 SPSS Macro 程序称为 MODMED 程序（modmed. sps），研究者可于下列网址下载使用：http://www.psych.ku.edu/preacher/w。兹将部分重点程序内容摘要如下：

```
/* 本程序的版权宣告 */.
/* MODMED version 1.1, posted January 23, 2007 */.
/* Written by Andrew F. Hayes */.
/* School of Communication */.
/* The Ohio State University */.
/* hayes.338@osu.edu */.
/* This revision adds bias corrected bootstrap confidence intervals to the
output */.
define MODMED (dv = ! charend('/')/med = ! charend('/')/dvmodel =
```

```
! charend('/')/mmodel = ! charend('/')
  /dvmodv = ! charend ('/') ! default (9999)/mmodv = ! charend ('/') ! default
(9999)/covmat = ! charend('/') ! default(0)
  /varord = ! charend('/') ! default(2)/boot = ! charend('/') ! default (1)).
/* 如欲取得 1st-order 方差，请改用 varord = ! charend('/') ! default(1) */.
preserve.
set seed = random.
set length = none.
set mxloop = 10000000.
/* This section does listwise deletion */.
save outfile = 'c:\modmed.sav'.
count miss = ! dv ! med ! dvmodel ! mmodel (sysmis).
select if (miss = 0).
save outfile = 'c:\mmtemp.sav'/drop = miss.
matrix.
……中间程序从略……
end matrix.
restore.
get file = 'c:\modmed.sav'.
! end define.
/* MODMED Macro 结束 */
/* 执行 MODMED 之控制叙述 */
MODMED DV = ATT/MED = POS/DVMODEL = POS NFC/ MMODEL =
MOOD NFC /BOOT = 1000.
```

前述最后一行控制指令是操作本程序的关键指令,研究者必须了解此行程序中的控制语法。兹将相关语法的涵义简单说明如下:DV 表依变量,MED 表中介变量,DVMODEL 表效标为依变量的模型,MMODEL 表效标为中介变量的模型,BOOT 表 bootstrap 样本数。至于"依变量模型"输入变量的顺序,依次为中介变量、调节变量(调节中介与依变量间的关系);"中介变量模型"输入变量的顺序,依次为自变量、调节变量(调节中介与自变量间的关系)。因此,由本行的控制指令可知,依变量为 ATT,中介变量为 POS,"依变量模型"输入变量的顺序,依次为中介变量 POS、调节变量 NFC,而"中介变量模型"输入变量的顺序,依次为自变量 MOOD、调节变量 NFC、Bootstrap 样本数 1000。执行此程序后,输出的报表如表 9-14 所示。

表 9-14　MODMED 的程序报表

```
Preacher, Rucker, & Hayes Moderated Mediation Analysis

You specified model number:
5
Variables in System:
    IV: MOOD
    DV: ATT
Med Var: POS
Mod Var: NFC
Sample size:
100
中介变量模型(MEDIATOR VARIABLE MODEL)
```

	Coeff	SE	t	P > |t|
Constant	.0404	.6971	.0579	.9540
MOOD	4.3357	.6971	6.2193	.0000
NFC	.7672	.5130	1.4956	.1380
Inter1	1.2565	.5130	2.4496	.0161

所得结果与表 9-12 中 SPSS 分析结果完全相同。

依变量模型(DEPENDENT VARIABLE MODEL)

	Coeff	SE	t	P > |t|
Constant	1.9394	1.2932	1.4996	.1371
MOOD	1.4802	1.5467	.9570	.3410
NFC	.3558	.9716	.3661	.7151
Inter1	-2.1691	1.0270	-2.1121	.0373
POS	1.2479	.1887	6.6130	.0000
Inter2	-.0359	.1291	-.2785	.7812

所得结果与表 9-13 中 SPSS 分析结果完全相同。

Interaction Terms:

Inter1: MOOD * NFC

Inter2: POS * NFC

调节型 Bootstrap 间接效果(-1SD, 0, +1SD)

Conditional indirect effect at specific value(s) of the moderator(s)

NFC	Boot Ind	Boot SE	Boot Z	Boot P
-1.4046	3.3640	1.6483	2.0409	.0413
.0000	5.3850	1.2377	4.3507	.0000
1.4046	7.1991	2.0406	3.5280	.0004

Moderator values listed are the sample mean and +/ -1 SD

本程序不仅提供调节变量值为 0 及上下一个标准差的 Bootstrap 间接效果:5.358,7.1991,3.364,而且提供所有调节变量值的 Bootstrap 间接效果。

调节型 Bootstrap 间接效果(全部)

Conditional indirect effect at range of values of the moderator(s)

NFC	Boot Ind	Boot SE	Boot Z	Boot P
-4.8265	-2.4254	4.8135	-.5039	.6143
-4.4314	-1.6942	4.2676	-.3970	.6914
-4.0363	-.9794	3.7777	-.2593	.7954
-3.6411	-.2810	3.3407	-.0841	.9330
-3.2460	.4011	2.9526	.1358	.8919
-2.8509	1.0668	2.6086	.4090	.6826
-2.4558	1.7161	2.3035	.7450	.4563
-2.0607	2.3491	2.0323	1.1559	.2477

NFC	Boot Ind	Boot SE	Boot Z	Boot P
-1.6656	2.9657	1.7913	1.6556	.0978
-1.2705	3.5659	1.5802	2.2567	.0240
-.8753	4.1498	1.4048	2.9539	.0031
-.4802	4.7173	1.2809	3.6827	.0002
-.0851	5.2684	1.2347	4.2671	.0000
.3100	5.8031	1.2937	4.4858	.0000
.7051	6.3215	1.4700	4.3002	.0000
1.1002	6.8235	1.7558	3.8862	.0001
1.4953	7.3091	2.1352	3.4232	.0006
1.8904	7.7784	2.5947	2.9978	.0027
2.2856	8.2313	3.1257	2.6334	.0085
2.6807	8.6678	3.7231	2.3281	.0199
3.0758	9.0880	4.3841	2.0729	.0382

由本表中的数据，可看出调节变量 NFC 的值越小，间接效果越小，其效果值为 0 的概率则越高；调节变量 NFC 的值越大，间接效果则越大，其效果值为 0 的概率则越低。

Indirect effect statistics are bootstrap estimates.
Bootstrap p-values assume normal bootstrap distribution.
Bootstrap 样本大小
本次统计分析使用了 1000 个 bootstrap 样本点。
Number of bootstrap samples：
1000
——END MATRIX——

（三）SEM 解法

利用 SEM 来分析条件间接效果，亦有两种方法：Amos 路径系数法与 Amos MCMC 法。首先谈第一种 Amos 路径系数法。图 9-46 是前述第五类型条件间接效果模型的路径设计，其条件间接效果的点估计（参见公式左侧）与一阶变异量（不包含画线部分）、二阶变异量（参见公式右侧）须利用下列公式计算（Preacher，Rucker & Hayes，2007），式中 w 表调节变量的值。

$$(\hat{a}_1 + \hat{a}_3 W) \times (\hat{b}_1 + \hat{b}_2 W) \qquad \begin{aligned} &(\hat{b}_1 + \hat{b}_2 W)^2 (s^2_{\hat{a}_1} + 2s_{\hat{a}_1 \cdot \hat{a}_3} W + s^2_{\hat{a}_3} W^2) \\ &+ (\hat{a}_1 + \hat{a}_3 W)^2 (s^2_{\hat{b}_1} + 2s_{\hat{b}_1 \cdot \hat{b}_2} W + s^2_{\hat{b}_2} W^2) \\ &\underline{+ (s^2_{\hat{a}_1} + 2s_{\hat{a}_1 \cdot \hat{a}_3} W + s^2_{\hat{a}_3} W^2)(s^2_{\hat{b}_1} + 2s_{\hat{b}_1 \cdot \hat{b}_2} W + s^2_{\hat{b}_2} W^2)} \end{aligned}$$

Amos 路径系数法，须利用上述间接效果的点估计与一阶变异量、二阶变异量（误差

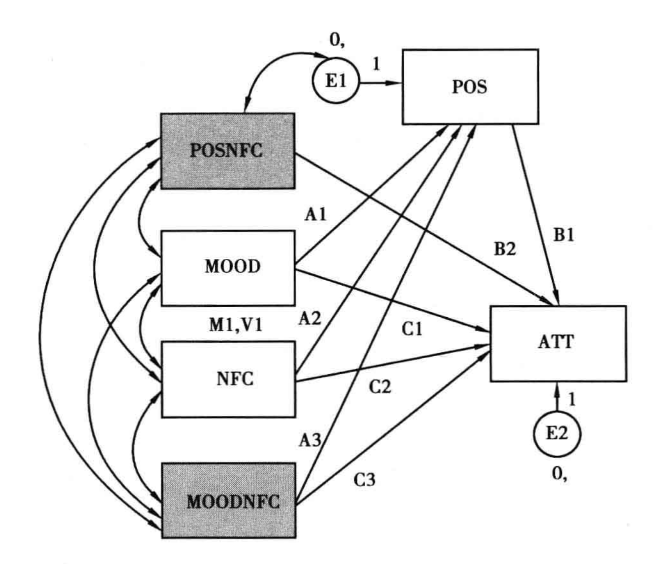


图 9-46 Amos 条件间接效果模型的路径设计

较小),计算 z 值,过程较烦琐,使用前述 MODMED 程序似乎较为简便。为了计算一阶变异量与二阶变异量,研究者必须在 Amos 分析属性窗口下的[Output]中,点选[Covariances of estimates],以取得相关的统计量。注意执行本模式时,须使用原始数据(使用离差分数)的平均数结构模型,在 Amos 分析属性窗口[Estimation]下,须点选[Estimate means and intercepts]。因使用离差分数表示,以致 MOOD 与 NFC 的平均数均为 0;其中调节变量 NFC 的方差为 1.9729(参见图 9-47),故其标准差为 1.405。为彰显原变量与交互作用项间的类别,特将交互作用项的外框线条加粗,请打开 Amos 对象属性窗口中的[Colors],针对[Text color]、[Line width]、[Fill type]等进行设定,参见图 9-48。

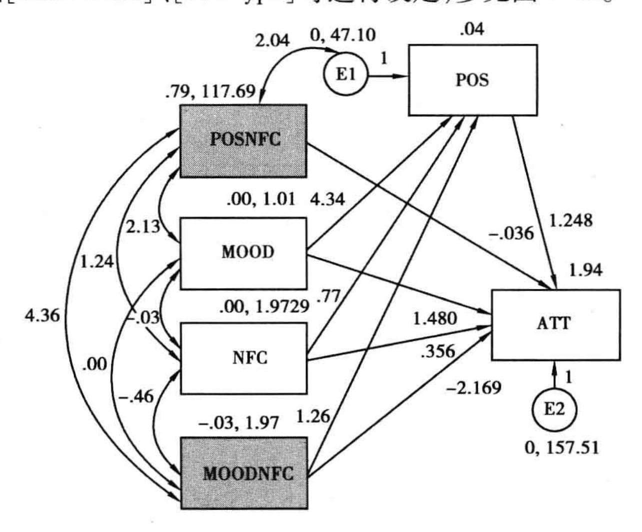


图 9-47 Amos 条件间接效果模型的路径系数

执行 Amos 路径系数法,首先,将图 9-47 中相关的数据,带入上述公式左侧的点估计公式中,以求得当调节变量 NFC 的值高于一个标准差 1.405(亦即 w = 0 + 1.405 = 1.405)时的间接效果。所得的间接总效果值等于:$(A1 + A3 * (M1 + \sqrt{V1})) \times (B1 + B2 *$

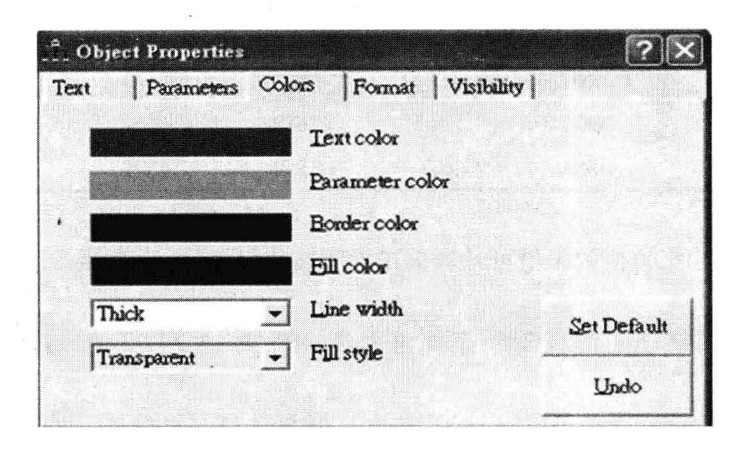


图 9-48　Amos 对象属性窗口中的[Colors]下路径图界面属性的设定

$(M1 + \sqrt{V1}) = (4.34 + 1.26 * 1.405) * (1.248 + -.036 * 1.405) = 7.317$。研究者配合图 9-46 中相关参数的标签，当能顺利求解。

其次，将图 9-49 中的对角线数据，带入上述的第五类型的公式右侧中，可求得其二阶变异量（NFC = +1SD），公式等于：$[1.248 + (-0.036)w]^2[0.4713 + 2(0.0019w) + 0.2551w^2] + (4.34 + 1.26w]^2[0.0338 + 2(-0.0007)w + 0.0158w^2] + [0.4713 + 2(0.0019)w + 0.2551w^2][0.0338 + 2(-0.0007)w + 0.0158w^2]$。研究者可配合图 9-46 中相关参数的标签，当能顺利求解。

最后，计算条件变异量，当 $w = +1.405$ 时（调节变量 NFC 的值高于一个标准差），带入上式可求得变异量：$(1.434 * 0.9802) + (37.336 * 0.06302) + (0.9802 * 0.06302) = 3.820$，其标准误为 $= \sqrt{3.820} = 1.955$，其间接总效果值：$(4.34 + 1.26 * 1.405)(1.248 + 1.405 * -0.036) = 7.317$。因此 $z = 7.317/1.955 = 3.74$，显示出当 $w = +1.405$ 时，间接总效果值达 0.05 显著水平。当 $w = -1.405$ 时（当调节变量 NFC 的值低于一个标准差），带入上式可得变异量：$(1.683 * 0.9695) + (6.603 * 0.06696) + (0.9695 * 0.066996) = 2.139$。其标准误为 $\sqrt{2.139} = 1.463$，其间接总效果值：$(4.34 + 1.26(-1.405))(1.248 + (-1.405 * (-0.036))) = 3.337$。因此 $z = 3.337/1.463 = 2.28$，显示出当 $w = -1.405$ 时，间接总效果值达 0.05 显著水平。

Amos Output

Variance-covariance Matrix of Estimates (Default model)

	C1	C2	C3	B1	B2	A1	A2	A3	V1	M1
C1	2.2715									
C2	.1773	.8964								
C3	.2662	.2784	1.0014							
B1	-.1451	-.0251	-.0408	.0338						
B2	-.0309	-.0192	-.0387	-.0007	.0158					
A1	.0000	.0000	.0000	.0000	.0000	.4713				
A2	.0000	.0000	.0000	.0000	.0000	.0082	.2551			
A3	.0000	.0000	.0000	.0000	.0000	.0019	.0595	.2551		
V1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0786	
M1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0199

图 9-49　Amos 估计参数的协方差矩阵

表 9-15 是 Amos 输出的直接效果报表。

表 9-15 Amos 直接效果报表

	POSNFC	MOODNFC	NFC	MOOD	POS
POS	.0000	1.2565	.7672	4.3357	.0000
ATT	-.0359	-2.1691	.3558	1.4802	1.2479

另外,由表 9-15 直接效果值可计算出排除间接效果之后,净直接效果(residual direct effect)的大小:

当调节变量 NFC 的值高于一个标准差 1.405 时:1.4802 - 2.1691 * 1.405 = -1.567;

当调节变量 NFC 的值低于一个标准差 1.405 时:1.4802 - 2.1691 * -1.405 = 4.527。

反映出净直接效果亦视 NFC 的高低而定。前述调节间接效果与净直接效果值,可绘制出交互作用图 9-50。

由图 9-50 可知,不管调节间接效果还是净直接效果,均须视 NFC 的高低而定。调节变量 NFC 的值越小,间接效果越小;调节变量 NFC 的值越大,间接效果则越大;而调节变量 NFC 的值越小,净直接效果越大;调节变量 NFC 的值越大,净直接效果则越小。

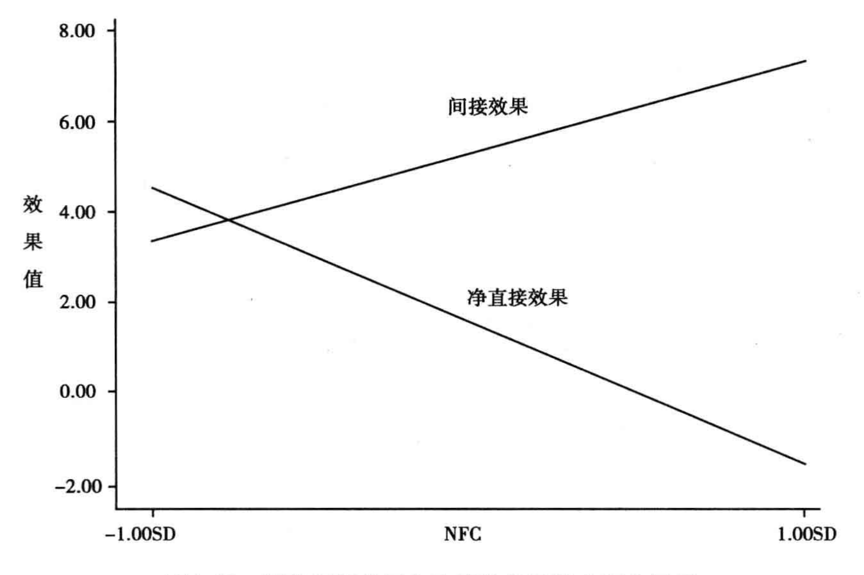


图 9-50 调节间接效果与净直接效果的交互作用图

其次,谈一谈第二种 Amos MCMC 法:本法须使用 Amos 的 Bayesian SEM,其路径图的设计亦如 Amos 路径系数法。并在 Bayesian SEM 窗口中,撰写计算条件型间接效果与自订参数的 MCMC 贝氏统计检验程序,程序内容摘要如下(文件名:MODMED. VB):

```
#Region "Header"
Imports System. MATH
Imports Microsoft. VisualBasic
Imports AmosEngineLib
Imports AmosEngineLib. AmosEngine. TMatrixID
```

```
Imports AmosExtensions. CustomEstimand
Imports PBayes
#End Region
Public Class CEstimand
    Implements IEstimand
    Public Sub DeclareEstimands( ) Implements IEstimand. DeclareEstimands
      newestimand( "SD" )
      newestimand( "W1 : -1" )
      newestimand( "W2:0" )
      newestimand( "W3 : +1" )
      newestimand( "Overall ModMed < 0" )
      newestimand( "Overall ModMed" )
      newestimand( " +1SD is less than zero" )
      newestimand( "MEAN is less than zero" )
      newestimand( " -1SD is less than zero" )
    End Sub
    Public Function CalculateEstimands( ByVal sem As AmosEngine) As
String Implements IEstimand. CalculateEstimands
      Estimand( "SD" ). value = SQRT( sem. ParameterValue( "v1" ) )
      'Conditional indirect effect at + -1SD Given a moderator
      ' Compute W1, W2, & W3, refer to PP. 858-859, Muller, Judd,
& Yzerbyt, 2005 for detail
      Estimand( "W1 : -1" ). value = ( sem. ParameterValue( "A1" ) + sem.
ParameterValue( "A3" ) * ( sem. ParameterValue( "M1" ) - Estimand( "SD" ).
value) ) * ( sem. ParameterValue( "B1" ) + sem. ParameterValue( "B2" ) * ( sem.
ParameterValue( "M1" ) - Estimand( "SD" ). value) )
      Estimand( "W2:0" ). value = ( sem. ParameterValue( "A1" ) + sem. Parame
terValue( "A3" ) * ( sem. ParameterValue( "M1" ) ) ) * ( sem. ParameterValue( "B1" ) +
sem. ParameterValue( "B2" ) * ( sem. ParameterValue( "M1" ) ) )
      Estimand( "W3 : +1" ). value = ( sem. ParameterValue( "A1" ) + sem.
ParameterValue( "A3" ) * ( sem. ParameterValue( "M1" ) + Estimand( "SD" ).
value) ) * ( sem. ParameterValue( "B1" ) + sem. ParameterValue( "B2" ) * ( sem.
ParameterValue( "M1" ) + Estimand( "SD" ). value) )
      'See Equation 7, Muller, Judd, &Yzerbyt, 2005 for detail
      'Compute total conditional indirect effect
      Estimand( "Overall ModMed" ). value = sem. ParameterValue( "B1" )
* sem. ParameterValue( "A3" ) + sem. ParameterValue( "B2" ) * sem.
ParameterValue( "A1" )
      Estimand( "Overall ModMed < 0" ). value = ( Estimand( "Overall
ModMed" ). value < 0)
      Estimand( " +1SD is less than zero" ). value = ( Estimand( "W3 : +1" ).
value < 0)
      Estimand( "MEAN is less than zero" ). value = ( Estimand( "W2:0" ).
value < 0)
      Estimand( " -1SD is less than zero" ). value = ( Estimand( "W1 : -1" ).
value < 0)
      Return ""      'Return an empty string if no error occurred
    End Function
End Class
```

注意程序中部分参数值的呼叫，例如，(sem. ParameterValue (" A1 ") 与 sem. ParameterValue(" A3") ，须与图 9-46 中的路径参数标签相一致，例如：A1 ~ A3，B1，B2。执行此程序后可得如图 9-51 窗口中的结果。

图 9-51 中 MCMC 数据显示，SD 表示 NFC 调节变量的标准差，其平均值为1.4754，W1：-1 表示低于调节变量平均数一个标准差，其 MCMC 平均间接效果为 3.2322；W2：0 表正好在调节变量的平均数，其 MCMC 平均间接效果为 5.4009；W3：+1 表高于调节变量平均数一个标准差，其 MCMC 平均间接效果为 7.3562。以上这些统计分析结果与前述 Bootstrapping 法的结果大致相同。根据 Muller, Judd 与 Yzerbyt（2005）公式 7 的定义（参见程序中定义），调节型间接总效果的平均值为 1.3984（0.95 置信区间：-0.02943 ~ 3.2912），在所有 MCMC 的样本中，其间接总效果小于 0 的概率为 0.0548；而当调节变量的数值高于调节变量平均数一个标准差时，在所有 MCMC 的样本中，其间接效果小于 0 的概率为 0.000，而当调节变量的数值低于调节变量平均数一个标准差时，在所有 MCMC 的样本中，其间接效果小于 0 的概率为 0.0111。因此，从条件型间接效果来看，NFC 值较高时，其间接效果较大 NFC 值较低时，其间接效果则较小。

Custom Estimands

File　Edit　View　Help

Numeric Estimands	Mean	S.D.	C.S.	95% Lower bound	95% Upper bound
SD	1.4754	0.1094	1.0006	1.2810	1.7028
W1:-1	3.2322	1.5679	1.0008	0.4501	6.6428
W2:0	5.4009	1.2591	1.0007	3.1032	8.0550
W3:+1	7.3562	2.1870	1.0006	3.4757	12.0796
Overall ModMed	1.3984	0.9125	1.0007	-0.2943	3.2912

Dichotomous Estimands	P	P1	P2	P3
Overall ModMed< 0	0.0548	0.0584	0.0551	0.0508
+1SD is less than zero	0.0000	0.0000	0.0000	0.0000
MEAN is less than zero	0.0000	0.0000	0.0000	0.0000
-1SD is less than zero	0.0111	0.0090	0.0140	0.0102

图 9-51　Amos Bayesian SEM 输出结果：自订参数

第六节　中介效果的实例解说

为让研究者进一步了解，如何运用 Amos 进行各种中介效果分析，特以吴丽华（2007）的 TSH 理论模型（参见图 9-52）做实例解说。以下将按①Sobel 间接效果的检验；②Amos 内建直接与间接效果检验；③Amos Bootstrap 内建直接与间接效果检验；④Amos 贝氏 SEM 自订参数估计法，逐一示范说明。

一、Sobel 间接效果法

由于“社会支持”对“教师工作压力”“教师工作压力”对“身心健康”等均具有显著直接效果，故可进一步检验“教师工作压力”在“社会支持”与“身心健康”之间的间接效果。“教师工作压力”在“社会支持”与“身心健康”之间是否具有显著间接效果，研究者可以利用前述 Sobel（1982）间接效果的检验公式进行检验（引自 Preacher & Leonardelli, 2006）：

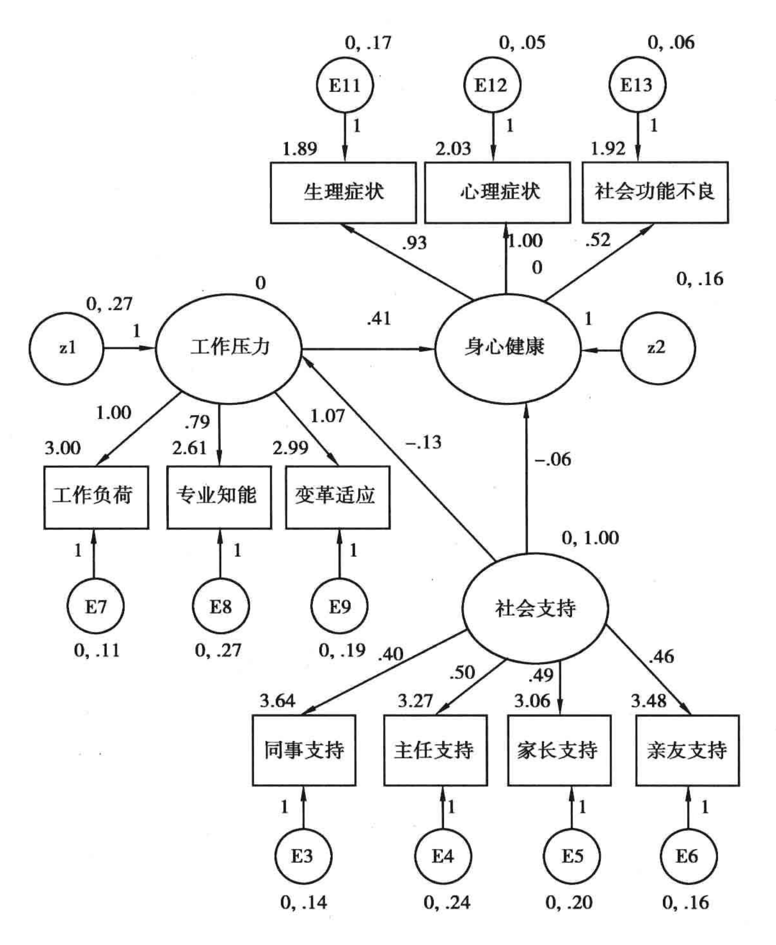


图 9-52　TSH 模型的路径设计图与未标准化参数估计值

$$Z = \frac{a \times b}{\sqrt{b^2 \times S_a^2 + a^2 \times S_b^2}}$$

公式说明：

a 为社会支持对教师工作压力的未标准化路径系数，b 为教师工作压力对身心健康的未标准化路径系数，S_a为社会支持对教师工作压力的未标准化路径系数的标准误，S_b为教师工作压力对身心健康的未标准化路径系数的标准误。

研究者由 Amos 报表中的相关回归系数可知，$a = -0.13$，$b = 0.41$，$S_a = 0.04$ 与 $S_b = 0.06$，代入上述间接效果的检验公式，求得 $Z = -2.94$，将 Z 值带入正态分布表，即可求得近似 p 值。本研究为求出较为精确的 p 值，乃将 $a = -0.13$，$b = 0.41$，$S_a = 0.04$ 与 $S_b = 0.06$等数值，输入能求出间接效果 p 值的软件中（Preacher & Leonardelli, 2006），求得 p 值为0.0033，参看图 9-53。

由于 $p = 0.0033 < \alpha = 0.05$，故“教师工作压力”在“社会支持”与“身心健康”之间具有显著的间接效果。由于社会支持对教师工作压力的未标准化路径系数为 -0.13，教师工作压力对身心健康的未标准化路径系数为 0.41，可求得社会支持对身心健康的“未标准化间接效果”为 -0.0535（ = -0.13 ×0.41）。此外，由于社会支持对教师工作压力的标准化路径系数为 -0.24，教师工作压力对身心健康的标准化路径系数为 0.47，可求得

社会支持对身心健康的"标准化间接效果"为 -0.11(= -0.24 ×0.47)。

	Input:		Test statistic:	p-value:
a	-.13	Sobel test:	-2.93495795	0.00333593
b	.41	Aroian test:	-2.90965915	0.00361823
s_a	.04	Goodman test:	-2.96092834	0.00306713
s_b	.06	Reset all	Calculate	

图 9-53　间接效果 p 值的网络计算软件

二、Amos 一般内建估计法

研究者可打开图 9-54 的分析属性窗口,点开[Output]窗口,勾选[Indirect, direct & total effects],让 Amos 自动计算总效果、直接与间接效果值的估计值。

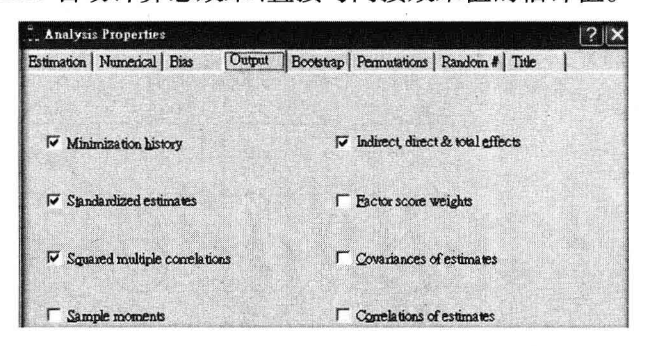


图 9-54　Amos 直接与间接效果值的设定

表 9-16 是 Amos 中有关总效果、直接与间接效果值的输出报表。总效果等于直接与间接效果值的和,例如:社会支持对于身心健康的总效果值为 -0.1183,等于直接与间接效果值的和(-0.0648 + -0.0535)。

表 9-16　Amos 总效果、直接与间接效果值输出报表

Total Effects(总效果)

	社会支持	工作压力	身心健康
工作压力	-.1302	.0000	.0000
身心健康	-.1183	.4114	.0000

Standardized Total Effects(标准化总效果)

	社会支持	工作压力	身心健康
工作压力	-.2447	.0000	.0000
身心健康	-.2521	.4660	.0000

Direct Effects(直接效果)

	社会支持	工作压力	身心健康
工作压力	-.1302	.0000	.0000
身心健康	-.0648	.4114	.0000

Standardized Direct Effects(标准化直接效果)

	社会支持	工作压力	身心健康
工作压力	-.2447	.0000	.0000
身心健康	-.1380	.4660	.0000

Indirect Effects(间接效果)

	社会支持	工作压力	身心健康
工作压力	.0000	.0000	.0000
身心健康	-.0535	.0000	.0000

Standardized Indirect Effects(标准化间接效果)

	社会支持	工作压力	身心健康
工作压力	.0000	.0000	.0000
身心健康	-.1140	.0000	.0000

由 Amos 报表可知,Amos 不会自动提供各直接与间接效果参数估计值的显著性检验,研究者如有需要可利用 Amos 内建 Bootstrap 法。

三、Amos 内建 Bootstrap 法

利用 Amos 内建 Bootstrap 方法,进行直接与间接效果的检验,不仅可以获得 p 值,还可获得置信区间。操作时,研究者须先打开分析属性窗口中的[Bootstrap]窗口(参见图9-55),勾选[Perform bootstrap]与[Percentile confidence intervals]或[Bias-corrected confidence intervals]。本例中,勾选了[Bias-corrected confidence intervals],bootstrap 样本数设定为 1000,区间的置信水平设定为 0.95。待分析的数据文件,如为协方差矩阵而非原始数据时,研究者必须点选[Monte Carlo](Parametric bootstrap)。

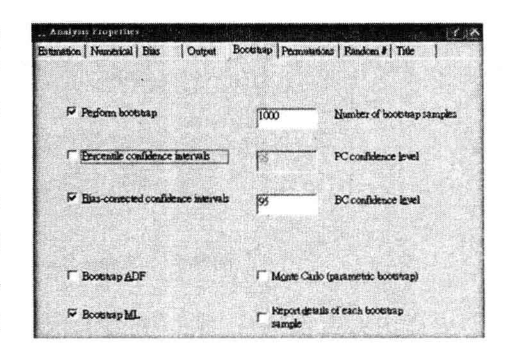


图 9-55 Amos Bootstrap 分析的设定

执行 Amos 之后,等出现图 9-56 的 Output 窗口后,在窗口左侧,点开[Estimates]下的[Matrices],就会出现[Total Effects]、[Direct Effect]、[Indirect Effect]等选目。当使用者

点选[Indirect Effect]选目时,即可点选左侧第二个小窗口中的[Bias-corrected confidence intervals],再点选[Twotailed Significance(BC)],窗口右侧就会出现"社会支持"对"身心健康"的间接效果(-0.0535)的显著性检验结果,p =0.015,意味着社会支持对身心健康的间接效果已达到统计上的0.05 显著水平。研究者如欲获得相关统计量的上下置信区间,可点选[Lower Bounds(BC)]与[Upper Bounds(BC)]。

当使用者点选[Direct Effect]选目时,图 9-57 窗口右侧就会出现社会支持对身心健康的直接效果(0.0648)检验结果,p=0.0659,意味着社会支持对身心健康的直接效果未达到统计上的0.05 显著水平;社会支持对工作压力的直接效果检验结果为 p=0.0209,意味着社会支持对工作压力的直接效果(0.1302)达到统计上的0.05 显著水平;而工作压力对身心健康的直接效果(0.4114)检验结果为 p=0.0020,意味着工作压力对身心健康的直接效果亦达到统计上的0.05 显著水平。

Indirect Effects - Two Tailed Significance (BC)

	社会支持	工作压力	身心健康
工作压力	...	...	...
身心健康	.0150	...	...
x6	...	...	...
x3	...	...	...
x4	...	...	...
x5	...	...	...
x7	.0209	...	...
x8	.0215	...	...
x9	.0236	...	...
y13	.0023	.0019	...
y12	.0025	.0020	...
y11	.0029	.0015	...

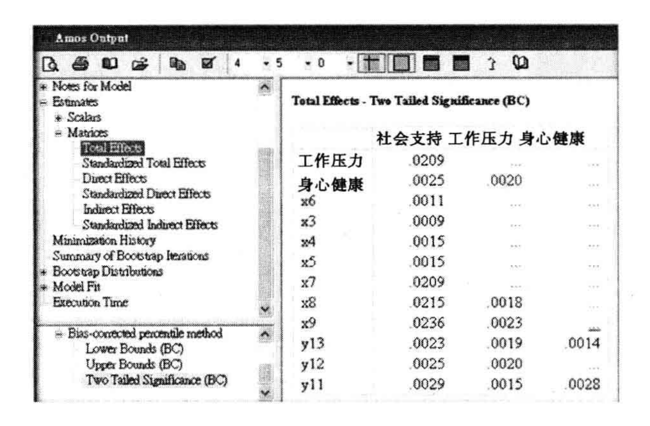
图 9-56 Amos Bootstrap 分析结果:间接效果检验

Direct Effects - Two Tailed Significance (BC)

	社会支持	工作压力	身心健康
工作压力	.0209	...	...
身心健康	.0659	.0020	...
x6	.0011	...	...
x3	.0009	...	...
x4	.0015	...	...
x5	.0015	...	...
x7	...	...	...
x8	...	.0018	...
x9	...	.0023	...
y13	...	...	.0014
y12	...	...	...
y11	...	...	.0028

图 9-57 Amos Bootstrap 分析结果:直接效果检验

当使用者点选[Total Effect]选目时,图 9-58 窗口右侧,就会出现社会支持对身心健康的总效果(-0.1183)检验结果为p =0 .0025,意味着社会支持对身心健康的总效果达到统计上的0.05 显著水平;其余社会支持对工作压力的总效果(-0.1302)检验结果为 p = 0.0209,意味着社会支持对工作压力的总效果达到统计上的0.05 显著水平;而工作压力对身心健康的总效果(0.4114)检验结果为 p=0.0020,意味着工作压力对身心健康的总效果亦达到统计上的0.05 显著水平。



Total Effects - Two Tailed Significance (BC)

	社会支持	工作压力	身心健康
工作压力	.0209	...	...
身心健康	.0025	.0020	...
x6	.0011	...	...
x3	.0009	...	...
x4	.0015	...	...
x5	.0015	...	...
x7	.0209	...	...
x8	.0215	.0018	...
x9	.0236	.0023	...
y13	.0023	.0019	.0014
y12	.0025	.0020	...
y11	.0029	.0015	.0028

图 9-58　Amos Bootstrap 分析结果：总效果检验

四、Amos 贝氏 SEM 自订参数估计法

使用贝氏 SEM 自订参数估计法，运用 MCMC 步骤研究者可以检验各种自订参数（参数交乘积、参数二次式等），用途广泛。Amos 操作步骤，须先画出如图 9-59 的提议模型（路径图中须标注路径参数 a，b，c）、连接数据文件、使用 Amos［Analyze］功能表单下的［Calculate Estimates］计算出模型中参数（a，b，c）的值。其次，使用［Analyze］功能表单下的［Bayesian Estimation］执行贝氏估计法，等 Bayesian SEM 表单上出现笑脸 后，按下暂停钮 暂停抽样，并按下 进行自订参数的估计，等 Amos 出现自订参数估计的程序设计窗口后，在 Amos 程序设计窗口中，撰写直接与间接效果估计值的相关程序（请参见以下程序设计：EX29. VB）。

```
#Region "Header"
Imports System
Imports Microsoft. VisualBasic
Imports AmosEngineLib
Imports AmosEngineLib. AmosEngine. TMatrixID
Imports AmosExtensions. CustomEstimand
Imports PBayes
#End Region
Public Class CEstimand
    Implements IEstimand
    Public Sub DeclareEstimands( ) Implements IEstimand. DeclareEstimands
      newestimand( "direct" )
      newestimand( "indirect" )
      newestimand( "difference" )
      newestimand( "indirect is less than zero" )
      newestimand( "indirect is smaller than direct" )
    End Sub

    Public Function CalculateEstimands( ByVal sem As AmosEngine) As
String Implements IEstimand. CalculateEstimands
      Estimand( "direct" ). value = sem. ParameterValue( "c" )
      Estimand( "indirect" ). value = sem. ParameterValue( "a" ) * sem.
ParameterValue( "b" )
```

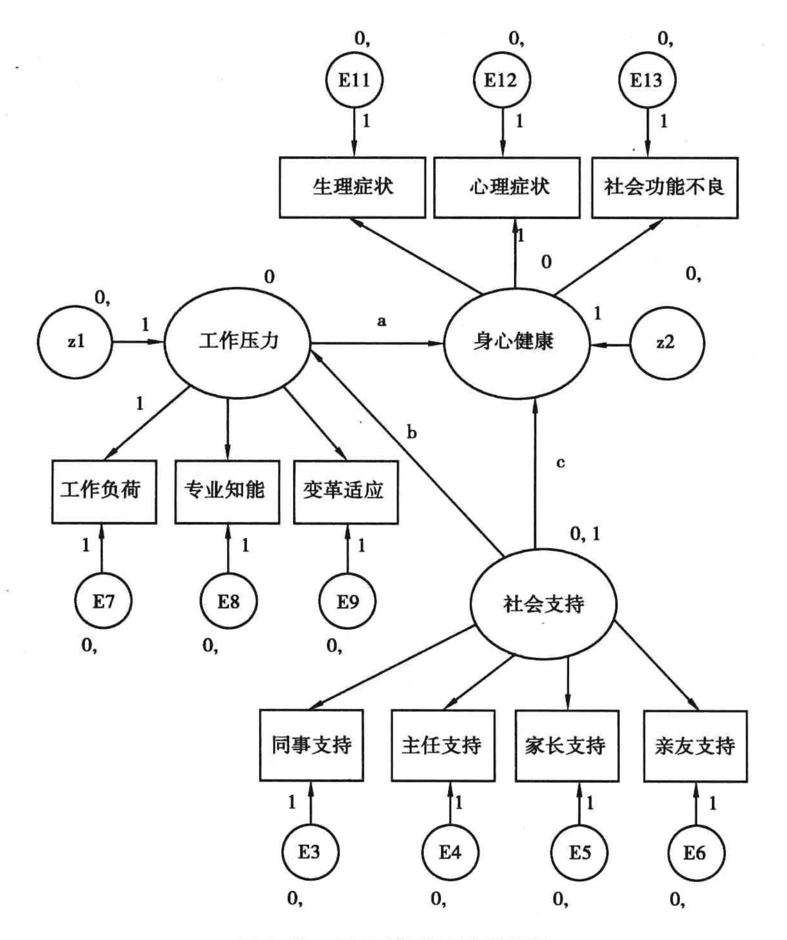


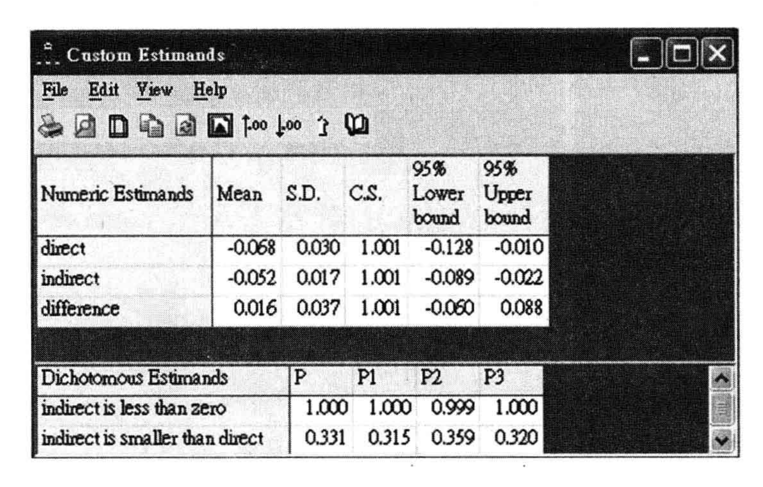
图 9-59 TSH 模型的路径设计

```
        Estimand("difference").value = Estimand("indirect").value -
Estimand("direct").value
        Estimand("indirect is less than zero").value = (Estimand("indirect").
value < 0)
        Estimand ( " indirect is smaller than direct " ) . value =
(Estimand("indirect").value < Estimand("direct").value)
        Return ""     'Return an empty string if no error occurred
    End Function

End Class
```

所得 MCMC 统计量:直接与间接效果的估计值及置信区间,请参看图 9-60。由图 9-60的数据知,间接效果的平均值为 -0.052,其0.95 置信区间为 -0.089 ~ -0.0022,就所有 MCMC 样本,此值小于0 的概率为1.0。因此,该间接效果等于0 的概率为0,达到统计上0.05 的显著水平;直接与间接效果差异的平均值为0.016,其0.95 置信区间为 -0.060 ~0.088(包含0),此值小于0 的概率为0.331,因此未达统计上0.05 的显著水平。

综合前述对于间接效果的三种统计分析结果,摘要于表 9-17。



Numeric Estimands	Mean	S.D.	C.S.	95% Lower bound	95% Upper bound
direct	-0.068	0.030	1.001	-0.128	-0.010
indirect	-0.052	0.017	1.001	-0.089	-0.022
difference	0.016	0.037	1.001	-0.060	0.088

Dichotomous Estimands	P	P1	P2	P3
indirect is less than zero	1.000	1.000	0.999	1.000
indirect is smaller than direct	0.331	0.315	0.359	0.320

图 9-60　MCMC 参数估计值

表 9-17　社会支持对于身心健康的间接效果摘要表

	估计值	P 值
Sobel 系数乘积法	-.0535	.0033
Bootstrap 法	-.0535	.0150
Bayesian SEM 法	-.0520	.0000

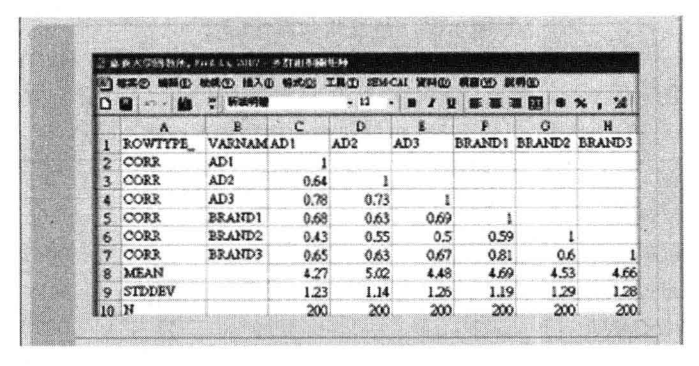
由表 9-17 的对照分析知，Sobel 系数乘积法、Bootstrap 法与贝氏估计法，所得间接效果的估计值与 P 值，三者间相差不大。

本章习题

1. 调节变量与中介变量的异同点何在？有哪些统计方法可以运用？

2. 何谓幽灵变量法（Phantom variable）？在 SEM 分析中，它有何用途？

3. 图 9-61 与图 9-62 的原始数据，是 MacKenzie 与 Spreng（1992）分析广告对于商标是否会因动机高低而有不同效应的数据。请利用 Amos 的多群组分析方法，检验与解释其中介效果。



	A	B	C	D	E	F	G	H
1	ROWTYPE_	VARNAM	AD1	AD2	AD3	BRAND1	BRAND2	BRAND3
2	CORR	AD1	1					
3	CORR	AD2	0.64	1				
4	CORR	AD3	0.78	0.73	1			
5	CORR	BRAND1	0.68	0.63	0.69	1		
6	CORR	BRAND2	0.43	0.55	0.5	0.59	1	
7	CORR	BRAND3	0.65	0.63	0.67	0.81	0.6	1
8	MEAN		4.27	5.02	4.48	4.69	4.53	4.66
9	STDDEV		1.23	1.14	1.26	1.19	1.29	1.28
10	N		200	200	200	200	200	200

图 9-61　低动机组原始数据

	A	B	C	D	E	F	G	H
1	ROWTYPE_	VARNAME_	AD1	AD2	AD3	BRAND1	BRAND2	BRAND3
2	CORR	AD1	1					
3	CORR	AD2	0.72	1				
4	CORR	AD3	0.76	0.74	1			
5	CORR	BRAND1	0.51	0.46	0.57	1		
6	CORR	BRAND2	0.32	0.33	0.39	0.4	1	
7	CORR	BRAND3	0.54	0.45	0.6	0.73	0.45	1
8	MEAN		4.35	4.93	4.59	4.86	4.71	4.74
9	STDDEV		1.24	1.18	1.2	1.1	1.08	1.15
10	N		160	160	160	160	160	160

图 9-62　高动机组原始数据

提示：首先，假设广告与商标建构的平均数及 Gamma 系数在组间均不相等，其余的组间参数均设定为相等，作为基础模型。其次，假设广告与商标建构的平均数及 Gamma 系数在组间均相等，作为第二个检验模型。第三个检验模型，广告与商标建构的平均数在组间相等，但 Gamma 系数在组间则设定为不相等。最后一个检验模型，广告与商标建构的平均数在组间设定为不相等，但 Gamma 系数在组间则设定为相等，参见图 9-63 与图 9-64 的路径设计。

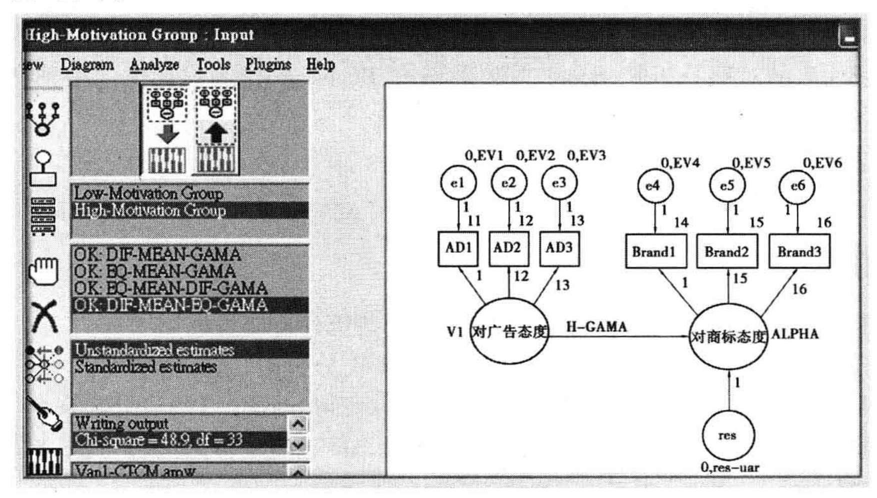


图 9-63　低动机组路径设计

4. 根据表 9-18 的 Amos 报表，分析多个模型比较的结果与意义。

表 9-18　Amos 高、低动机群组理论模型比较分析

Model	NPAR	CMIN	DF	P	CMIN/DF
DIF-MEAN-GAMA	22	41.2995	32	.1256	1.2906
EQ-MEAN-GAMA	19	50.3752	35	.0447	1.4393
EQ-MEAN-DIF-GAMA	20	42.8723	34	.1415	1.2609
DIF-MEAN-EQ-GAMA	21	48.8726	33	.0370	1.4810
Saturated model	54	.0000	0		
Independence model	24	1381.8683	30	.0000	46.0623

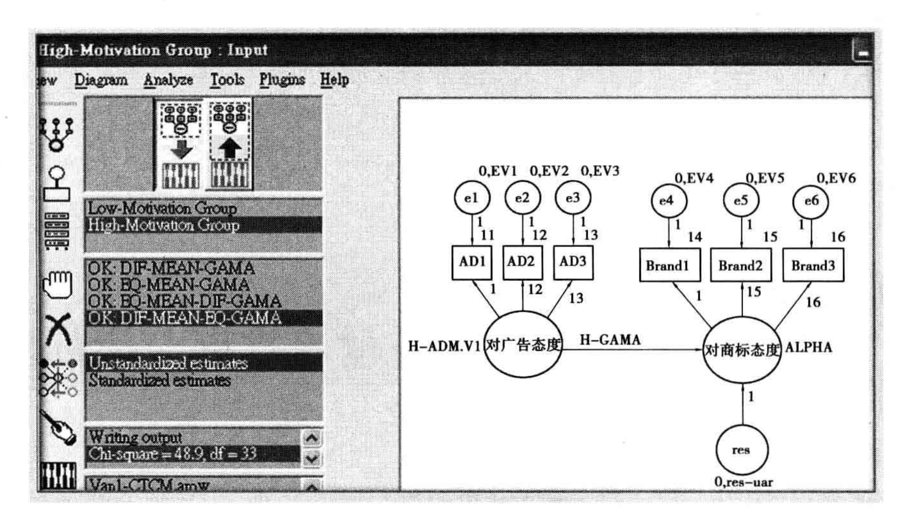


图 9-64　高动机组路径设计

第十章　测量工具的编制:量表发展与指标建构

量表发展(Scale development)涉及效果指标(Effect indicators)的编制,而指标建构(Index construction)涉及原因指标(Causal indicators)的编制。这两类测量工具的发展理念与过程并不全然相同,但在研究文献上出现不少的迷思。根据 Jarvis,Mackenzie 与 Podsakoff(2003)调查 1977—2000 年间四种著名营销杂志(*The Journal of Consumer Research*, *Journal of Marketing*, *Journal of Marketing Research*, *Marketing Science*)中有关 SEM 的论文,发现有 29% 的论文把原因指标或效果指标与建构间的因果关系弄反了。同样地,Petter,Straub 与 Rai(2007)调查 2003—2005 年间两种知名杂志 *MIS Quarterly* 与 *Information System Research* 杂志中的论文,发现有 30% 的论文把原因指标的建构误为形成性指标的建构;Freeze 与 Raschke(2007)调查 2003—2006 年间知名杂志 *MIS Quarterly* 杂志中的论文,发现有 71% 的论文未说明文中所运用的建构是形成性的还是反映性的。此种测量模型界定错误或根本漠视不理的迷思,值得其他领域研究者警惕,在测量工具的发展前就需要先厘清您的构念到底是原因/形成性指标(Causal/formative indicators)的建构,还是效果/反映性指标(Effect/reflective indicators)的建构。此种测量模型的界定错误(Measurement model misspecification)除了会影响到测量工具的内容效度之外,亦可能会导致结构关系的错误结论(Jarvis, Mackenzie, & Podsakoff, 2003)或结构路径系数的偏估(Law & Wong, 1999),及导致偏高的第一类型错误(会发生在路径方向错误的效果指标)或偏低的第二类型错误(会发生在路径方向错误的原因指标)(Mackenzie,Podsakoff & Jarvis, 2005; Freeze & Raschke, 2007; Petter, Straub & Rai, 2007)。因此,研究者在建构 SEM 理论模型时,必须先根据实质理论与研究目的将原因指标或效果指标与建构间的因果关系厘清,才能正确编制出具有信度、效度的研究工具。

第一节　反映性指标与形成性指标的定义

依照 Jarvis,Mackenzie 与 Podsakoff(2003)的定义,原因指标应符合以下七大条件:①建构的特征依靠指标的界定;②指标的改变会造成建构的改变;③建构的改变不会造成指标的改变;④指标不一定要归属于共同主题之下;⑤某一指标值的改变并不一定会造成其他指标值的改变;⑥删除一个指标可能会导致建构范畴的更动;⑦指标不需要有相同的前提与后果。因而原因或形成性指标,是指观察指标乃是潜在变量的因,涉及此类工具的发展常被称为“指标建构”,常见于经济学、市场营销与管理学;而效果或反映性

指标,系指观察指标乃是反映潜在变量的效果指标,涉及此类工具的发展常被称为“量表发展”,最常见于教育学或心理学。

换言的,当一个心理建构是外显指标的果时,这些外显指标就称为原因指标(Bollen, 1989; Jarvis, Mackenzie & Podsakoff, 2003; MacKenzie, Podsakoff & Jarvis, 2005)。例如图 10-1 中,造成工作满意的原因主要来自于:职务的满意、薪水的满意、对同事的满意、对上司的满意、升迁的满意。这五个满意外显指标即为工作满意的原因指标(MacKenzie, Podsakoff & Jarvis, 2005)。这五个指标并不在于测量相同的潜在构念。这个“工作满意”建构的属性,显然是一个多维度的,而这个实征建构由这些外显指标的总变异量所界定。因此,外显指标间不一定具有高相关,请注意在图 10-1 中,各指标间的共变关系路径均带有 ? 号,此类建构的测量特质显然与传统的单维度建构基本要求不同。利用原因指标建构出来的测量工具,在研究文献上有时又称为指标(An index),而利用效果指标建构出来的测量则称为量表(A scale)。因而在测量工具的发展过程当中,前者称为指标建构,后者称为量表发展。就单一向度而言,效果指标重视指标间的内部一致性,而原因指标的内部一致性则不重要,重要的是不能漏掉任何重要的原因指标。因此,传统的信、效度概念(如 Cronbach α)不适用在这类指标建构的编制上(Bollen & Lennox, 1991; Jarvis, Mackenzie & Podsakoff, 2003; MacKenzie, Podsakoff & Jarvis, 2005; Diamantopoulos & Siguaw,2006)。

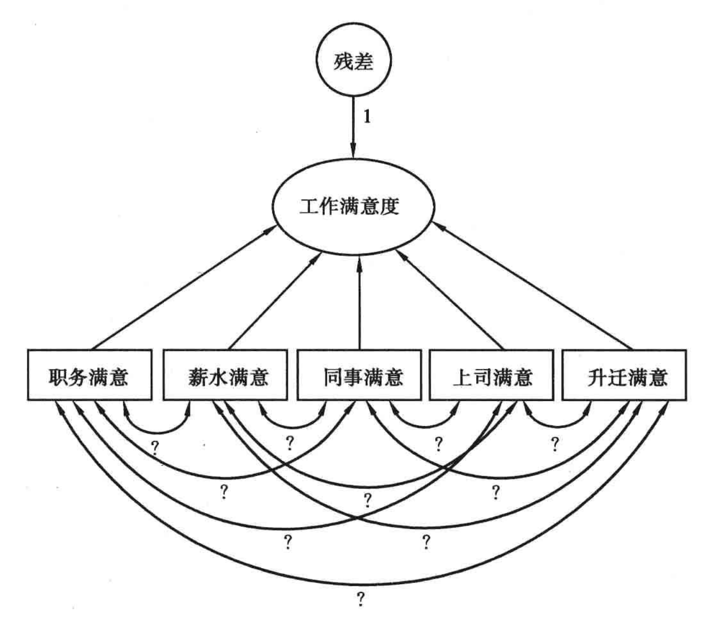


图 10-1　原因指标路径图

一般来说,研究者都希望预测变量与效标具有密切相关,但预测变量间则不可具有高度相关。建构一个有效的原因指标,基本上偏重广度,即不要把重要的预测变量遗漏了,但也不要把不重要的预测变量纳进来,否则建构层次的残差会很大。由此观之,当此残差很大时,可能就是反映出该建构的某些重要指标被遗漏了。当一个心理建构被视为外显指标的因时,这些外显指标就称为效果指标或反映指标,亦即心理建构为外衍变量,而外显指标为内衍变量(Bollen, 1989),而这个实征建构由这些外显指标的共同变异量

所界定。一个外显指标到底是因是果还是两者皆是,乃是研究者确定因果模型前必须先考虑的问题,因为这项决定会改变观察指标的选择(Petter, Straub & Rai, 2007)。

第二节　反映性指标与形成性指标的特点与不同

反映性指标与形成性指标的编制,在建构的界定与拟题阶段上,两者并无差异,但因两者所使用的项目分析方法不同,最后选题的结果就会有很大差异,测量编制者不能等闲视之,以免发展出来一个统计上正确但理论上却是错误的量表。以下将从相关文献中摘要出几个重要判定标准与相异点(Diamantopoulos,1999; Jarvis, Mackenzie & Podsakoff, 2003; Freeze & Raschke, 2007; Wilson,Callaghan & Stainforth, 2007; Coltman, Devinney, Midgley & Venaik, 2008),以利研究者分辨反映性指标测量模型(reflective indicator model)与形成性指标测量模型(formative indicator model),进而编制出理想的测量工具。这两种测量模型,其数学定义是完全不同的。反映性指标与共同潜在变量(common latent variable)间的数学定义如下:

$$Y_i = \beta_{i1}X_1 + \varepsilon_i$$

式中　Y_i = 第 i 个指标

β_{i1} = 路径系数

X_1 = 潜在变量

ε_i = 指标 i 的测量误差

至于形成性指标与组合型潜在变量(composite latent variable)间的回归关系,则定义如下:

$$Y = \beta_1X_1 + \cdots + \beta_nX_n + \zeta$$

式中　Y = 组合变量

β_i = 回归系数

X_i = 指标分数

ζ = 残差

这两种因果方向不同的测量模型,可具体界定如图 10-2:

据此路径模型,将两个测量模型的不同之处摘要说明如下:

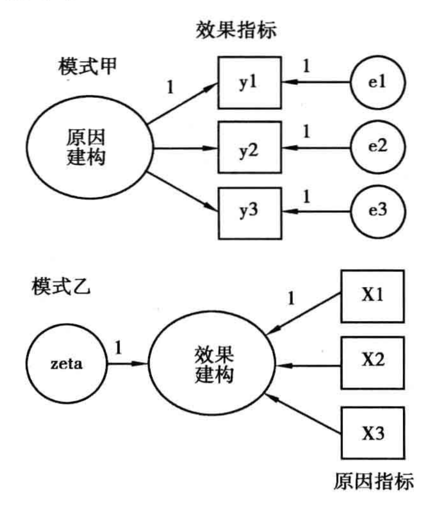


图 10-2　反映性模型(甲)与形成性模型(乙)

一、因果关系相反

由图 10-2 知,反映性模型的因果关系是从建构到指标(模型甲),而形成性模型的因果关系是从指标到建构(模型乙),两者的因果关系完全相反。因此,反映性测量模型中的指标是建构的外显表征,指标的改变不会造成建构意义的改变,但建构意义的改变会造成指标的改变,因为指标为建构所决定;而形成性测量模型中的指标乃是用来界定建构的特质,指标的改变会造成建构意义的改变,因为建构为指标所组成,但建构意义的改变并不必然造成指标的改变(Bollen & Lennnox, 1991)。

二、内部一致性要求不同

反映性模型中的指标间须具有共变关系,而形成性模型的指标间不一定具有共变关系。因此,反映性模型强调共同变异量,该建构的指标须具有内部一致性(internal consistency)及信度;而形成性模型强调独特变异量,该建构的指标间并不一定具有内部一致性,因而通常无法进行指标间的信度评估。

三、指标特质的互换性

反映性模型中的指标因分享共同主题或内涵而具有互换的性质,指标的增减对于该建构的概念范畴或主要特质影响不大;而形成性模型中的指标因未必分享共同主题或内涵而不具有互换的性质,指标的增减会严重影响该建构的概念范畴。

四、测量误差的层次不同

反映性模型中的测量误差在指标层次,可通过共同因素分析将测量误差抽离出来;而形成模型中的误差项在建构层次(比较图 10-2 中甲与乙的相异处),其指标假定无测量误差,指标品质的评估可通过消失四分项(vanishing tetrad test)的统计分析(详如后续说明)与指标多元共线性的检查。

五、建构指标的前因与后果的一致性

反映性模型中的测量指标与相关的建构间须具有相同的前因与后果,因此,量表效度的评估可以通过传统的内容效度、聚敛效度与区辨效度加以检验;而形成性模型中的测量指标与相关的建构间不一定具有相同的前因与后果(参见图 10-26 中(b)和(c)),因此,量表效度的评估须通过 MIMIC(multiple indicators and multiple causes)之类的模型进行构念网络效度(nomological validity)分析。

六、建构的向度

当使用形成性指标来诠释一个建构的各个独特面向时,该建构一定是多向度(multi-dimensionality)的概念(Wilson, Callaghan & Stainforth, 2007)。当指标所属的建构仅具单一向度(uni-dimensionality)时,该建构肯定是反映性建构;假如指标所属的建构是多向度的,那么研究者就须根据建构的实质内涵与因果特性,判别它的属性:反映性或形成性。

研究者根据该建构的相关理论与前述的异同点,进行逻辑比较分析(a logic check)之后,就比较容易分辨建构与指标间的关系与方向。如果尚有疑虑,亦可进行四个变量为一组的消失四分项(vanishing tetrad)协方差统计分析(Bollen & Ting, 2000),以决定指标及建构的测量方向性。反映性或形成性,这是一种验证性的统计分析(Wilson, Callaghan & Stainforth, 2007)。消失四分项协方差统计分析与传统 SEM 略有不同,分析时须以四个变量为一组进行消失四分项检验,例如,设有四个变量:g, h, i, j,则共有六个协方差,可以产生三组四分项:τ_{ghij}、τ_{gijh}、τ_{gjhi}。其中 τ_{ghij} 协方差积差统计量的定义如下:$\tau_{ghij}=\sigma_{gh}\sigma_{ij}-\sigma_{gi}\sigma_{hj}$。在效果指标的模型下,式中第一个交乘积项 $\sigma_{gh}=\lambda_g\lambda_h\varphi$, $\sigma_{ij}=\lambda_i\lambda_j\varphi$,$\sigma_{gh}\sigma_{ij}=\lambda_g\lambda_h\lambda_i\lambda_j\varphi^2$($\varphi$ 表示潜在变量的方差),而式中第二个交乘积项$\sigma_{gi}=\lambda_g\lambda_i\varphi$, $\sigma_{hj}=\lambda_h\lambda_j\varphi$, $\sigma_{gi}\sigma_{hj}=\lambda_g\lambda_i\lambda_h\lambda_j\varphi^2$,因此,$\tau_{ghij}=0$。假如 $\tau_{ghij}=0$,该四分项就称为"消失四分

项”,效果指标(effect-indicator)模型下(即形成性建构)的四分项应皆为消失四分项(至于为什么会等于0,细节请参看 Bollen & Ting, 2000 ,p. 7)。此种 SEM 检验法,不仅可正式评估模型的适配度(传统的适配度指标大都无法进行正式的统计检验),且不需要估计参数,因此费时较少。而且,此种消失四分项检验方法,尚可评估一些不可辨识的模型,及检验传统 SEM 无法检验的非嵌套模型(但须 tetrad-nested:一个模型的隐含四分项为另一模型四分项的次集合,参见文末习题8的例子)。消失四分项协方差统计分析有四个基本步骤(Hipp, Bauer & Bollen, 2005):

1. 界定指标与建构的理论模型,以便取得样本协方差矩阵与隐含协方差矩阵;
2. 决定模型隐含的消失四分项;
3. 去除重复的消失四分项,仅保留非线性相依的四分项;
4. 所有独立消失四分项同时检验。

依据定义,在反映性测量模型中,此种协方差乘积的差异量 τ_{ghij},τ_{gijh},τ_{gjhi}应等于0,但在形成性测量模型中,则通常并非如此,而会出现显著性的检验结果。因此,拒绝虚无假设 H_0: $\tau=0$时(例如:$p < 0.05$),意指形成性指标测量模型可能为真;而无法拒绝时(例如:$p > 0.05$),意指反映性指标测量模型可能为真,亦即反映性指标所隐含的非重复四分项的值接近于0。不过,研究者仍须进行以下两种验证:第一、指标间的相关系数是否显著相关;第二、进行反映性指标模型的检验,确认是否大部分的路径系数或潜在变量的变异量是否显著地大于0,不能通过这两种检验时,形成性指标的测量模型较可能为真(Bollen & Ting, 2000),第三、样本大小不可过大或过小,否则不能正确决定指标的特质。当变量数较多时,本统计检验过程非常繁复,Hipp,Bauer 与 Bollen(2005)乃设计了一个 SAS 模块:“CTANEST1”,大大地简化了四分项协方差统计分析的时程,其操作步骤请参见本书附录六说明。有需要的读者,可以前往他们的网站下载该 SAS 模块。至于应用实例解说,请参阅 Hadaya, Lege 与 Croteau (2005); Glanville 与 Paxton (2007); Wilson, Callaghan 与 Stainforth(2007)的论文。

第三节 形成性测量的争议

形成性测量(formative measurement)容易导致构念在不同模型上解释的困惑(interpretational confounding), Howell, Breivik 与 Wilcox (2007a) 在心理学方法期刊(*Psychological Methods*)上,首先主张研究者最好使用反映性测量(reflective measurement),因而引发了此一议题的论战。(Burt, 1976)定义解释上的困惑为:研究者在参数估计后,发现所赋予形成性建构的实征意涵(empirical meaning)与该建构的原先定义的名义意涵(nominal meaning)出现分歧,而导致对于该建构的推论与解读产生疑义。事实上,此种解释分歧的现象不仅出现在 SEM 上,也常出现在因素分析的解释上。例如,过去因素分析的使用者,就常发现相同的数据在不同的结构模型上,建构相关的因素负荷量却出现显著差异,此即该因素负荷量不具有测量不变性(measurement invariance),此时研究者的首要任务就是改善测量模型。在 SEM 分析中,假如一个建构的参数估计值(如因素负荷量、测量误差与建构的变异量)会因建构间路径的增删而显著改变,此即解释性困惑,因为结构模型的改变影响了该建构的实征意义。由此观之,解释

上的困惑关心的不是建构名称是否相同,而在于模型参数估计后,其实征意义是否改变。

Bollen(2007)在同期期刊上则撰文提出反驳,认为出现解释性困惑并不是形成性指标本质上的限制所致,而是模型界定错误(如遗漏了重要的路径与变量,或包含了不必要的路径与变量)所致。假如模型界定正确,不管研究者是使用形成性指标还是反映性指标,都不会产生解释性困惑。Bagozzi(2007)在同期期刊上亦加入论战,除了有条件地同意 Howell,Breivik 与 Wilcox(2007a)的论点的外,亦认为系统偏差(如共同方法偏差)亦可能是释性困惑的来源。接着,Howell,Breivik 与 Wilcox(2007b)则针对 Bollen 与 Bagozzi 的论点,一一加以澄清与辩驳,都认为自己的主张仍是有效的。欲知上述精采论战中几个虚拟实例详细内容的读者,请参阅英文原著或黄芳铭与李俊贤(2008)的文章。整合学者们的论点,笔者有以下几点看法与心得:

1. 除了模型界定错误、系统偏差会形成解释性困惑之外,组合变量本身是否单向度亦是困惑的来源。因为形成性模型的测量指标与相关的建构间不一定须具有相同的前因与后果,显示出此建构并非单一特质的建构,硬将它组成一个潜在变量,会导致建构意涵的混淆。
2. 形成性指标的建构,本质上是一个多向度的建构,是由一组固定的相互独立的指标组成,研究者硬将它们组在一块儿,委实有点不合乎逻辑,而当外部效标间相关性不高时,也易缺乏外部一致性(external consistency),这是形成性测量的一个困境,急待后续的深入研究。外部一致性的实例说明,请参看本章习题 10。
3. 研究者比较容易检查出反映性测量模型是否有解释性困惑,因此,可能的话研究者应尽可能少用形成性测量模型。因为建构本身并无反映性或形成性之分,其具有的属性会因测量的方法或内容而改变。换言之,虽然有些建构本质上是形成性的,但在概念层次上,大部分的建构均可以通过反映性测量测得。
4. 诚如 Bagozzi(2007)所言,形成性测量的运用会受许多本质上的限制。例如,单独存在的模型不可辨识、其参数无法估计、信效度不易建立。因而,在使用的范畴上会比反映性测量来得狭隘。

第四节　量表发展的过程

量表发展(Scale development)意指反映性指标的编制,强调题目间的相关性,其共同变异量、内部一致性与单向度的信度和效度检验过去学界已有良好的准则可资运用。测量编制者一般都遵循量表发展的五个基本原则(Bollen & Lennox, 1991):

1. 同一建构的外显指标应具有高度内部一致性;
2. 同一建构的外显指标间相关越高越好;
3. 单一维度建构中,信度相同的外显指标,本质上是可以替换的;
4. 同一建构内的相关应高于建构间的相关;
5. 指标的线性组合可以取代潜在变量。

至于此种量表发展的具体编制步骤,编制者一般须考虑以下几个面向:

一、内容面向的界定

内容面向的界定(Content specification)旨在构念向度的厘清,“全面彻底的文献探讨+请教该领域专家+实地观察”是唯一可行的路。单一向度(unidimensionality)的检验是必要工作,常用的统计方法是探索或验证性因子分析。

二、观察指标的界定

观察指标的界定(Indicator specification)与内容面向的界定具有密不可分的关系。观察指标的选择须顾及指标的广泛性(comprehensive)与周延性(inclusive),亦即观察指标的选择要能涵盖该建构的所有面向,且题目间要具有一定的重复性,以增强内部一致性,亦即广度与深度须兼具。

三、指标的量化

指标量化的主要任务在于决定使用什么量尺(the scale),以分派数字给不同层次的指标。社会科学最常见的是 Likert-type 五点量尺,自然科学最常见的是摄氏或华氏温度计。

四、指标间相关性的检验(indicator collinearity)与信度检验

内部一致性分析(internal consistency)与信度检验是传统测量编制者强调的重点工作,希望能借由它来检验自己编制的测量是否具有同质性。一般来说,研究者都会使用 Cronbach α 系数或计算题目与总分的相关,来分析选题。在 SEM 的分析体系下,可利用题目的信度指标(item reliability)为 R^2,此信度指标即为各标准化路径系数的平方值,作为选题的依据,一般都希望能大于 0.50。在选题的过程中,须特别注意题目内容的适切性(adequacy of a measure),不要一味追求统计特质的完美而忽视内容的适切性。换言之,利用统计指标盲目地删除题目以追求一个同质性的测量,可能毁及测量的内容效度。

此外,根据 Bollen(1989, p.216)与 Graham(2006)的理论说明,Cronbach α 系数植基于平行测量(单向度、负荷量相等、误差方差相等)或 tau 等值测量(单向度、负荷量相等、误差方差不等)的基本假设,如有违反则易导致低估内部一致性信度,尤其当测量的题数不多、题目的标准差变异大及题目间使用不同量尺时。一般来说,平行测量的基本假设相当严苛,要求每一指标来自于因素建构的负荷量与误差变异量均相等,很难达到这些基本要求。Graham(2006)因而主张在计算 Cronbach α 系数之前,先利用 SEM 检验测量模型,以进行相关基本假设的检验。如果发现该测量模型仅能符合同质性(congeneric)测量(除单向度之外、负荷量与方差均不等)的基本假设,研究者最好使用 SEM 来估计 Cronbach α 系数或使用建构信度(Construct/Composite reliability),以免低估内部一致性信度。图 10-3 是一个数常识测量,包括四个分测量(form1 ~ form4)的描述统计与相关矩阵,其中 form1 ~ form4 的量尺为八点量尺,form5 是 form4 量尺转换后的四点量尺分数。以下将以此数据比较说明 Cronbach α 在违反平行测量假设下的低估现象。

内部一致性,事实上与相关系数有密切关系,研究者须知道标准化的 Cronbach α 系数,适用于各题目的变异量差异不大时,或使用各题的标准分数计算总分时(Cortina, 1993)。因此,标准化 Cronbach α 系数,就可利用相关系数估计,其计算公式如下:

$$\rho_\alpha = \left[\frac{k}{k-1}\right]\left[\frac{2\sum_{i=1}^{n}\sum_{j\neq i}^{n} r_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{n} r_{ij} + 2\sum_{i=1}^{n}\sum_{j\neq i}^{n} r_{ij}}\right]$$

$$k = \left[\frac{k\bar{r}}{1+(k-1)\bar{r}}\right]$$

alpha-sem-cor.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

15 : form4

	rowtype_	varname_	form1	form2	form3	form4	form5
1	n		1248.000	1248.000	1248.000	1248.000	1248.000
2	corr	form1	1.000	.160	.278	.243	.241
3	corr	form2	.160	1.000	.175	.161	.159
4	corr	form3	.278	.175	1.000	.195	.192
5	corr	form4	.243	.161	.195	1.000	.998
6	corr	form5	.241	.159	.192	.998	1.000
7	stddev		2.821	2.794	3.100	2.881	1.244
8	mean		3.890	4.190	4.060	4.250	1.830

图 10-3　数常识的 SPSS 相关矩阵

利用这四个题目(form1 ~ form4)的测量的相关矩阵,带入上述的公式可求得 Cronbach α 系数为 0.503。根据原始数据,利用 SPSS 的 Reliability 副程序求得的 Cronbach α 系数亦为 0.503,因为 form1 ~ form4 的标准差甚为接近,其未标准化与标准化的 Cronbach α 系数均相同。

表 10-1　SPSS Cronbach α 系数报表

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.503	.503	4

当研究者建立如图 10-3 的数据文件后,就可根据平行测量的基本假设绘制如图10-4的路径图,因为负荷量与误差变异量均相等,在此路径图中相关路径系数均标注为相同的名称(w1 & v1)。图中 T 建构代表真分数的潜在特质,用以估计真分数的变异量,X 代表观察分数的组合变量(路径系数全部设定为 1),用以估计观察分数的变异量。在 SEM 的分析架构中,欲估计 Cronbach α 系数,只要将 T 与 X 的隐含相关系数平方即可(Graham, 2006)。欲知细节请参见 Allen 与 Yen(1979)与 Miller(1995)的信度系数(ρ_{XT}^2)定义:$\rho_{XT}^2 = \left(\frac{\sigma_{XT}}{\sigma_x\sigma_T}\right) = \frac{\sigma_T^2}{\sigma^2 X}$,亦即信度是观察分数变异量与

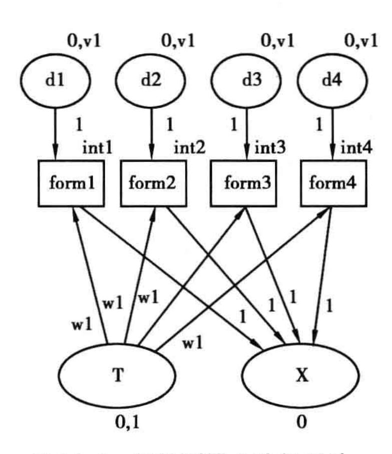


图 10-4　平行测量之路径设计

真分数变异量的比值。这是传统测量理论的主要依据，详细公式推演过程请参见 Allen 与 Yen(1979)的专著第 62 及 66 页公式 4 的说明。

由图 10-5 Tau 等值测量模型的参数限制可知，四个路径系数均标注为相同名称(w1)，代表欲进行路径系数等同的限制，而图 10-6 同质性测量模型的参数均未限制，可知负荷量与误差变异量均开放估计。前述三个测量模型分开设计与估计，稍嫌麻烦。研究者亦可更简便地设计在同一路径图中(参见图 10-7)，再利用 Amos 的模型管理窗口，进行模型中参数等同限制。此种分析途径除了省时省力之外，尚可进行三个模型间的适配度统计检验，可说一举数得。具体的操作步骤说明如下：首先建立如图 10-7 SEM 的信度分析模型。

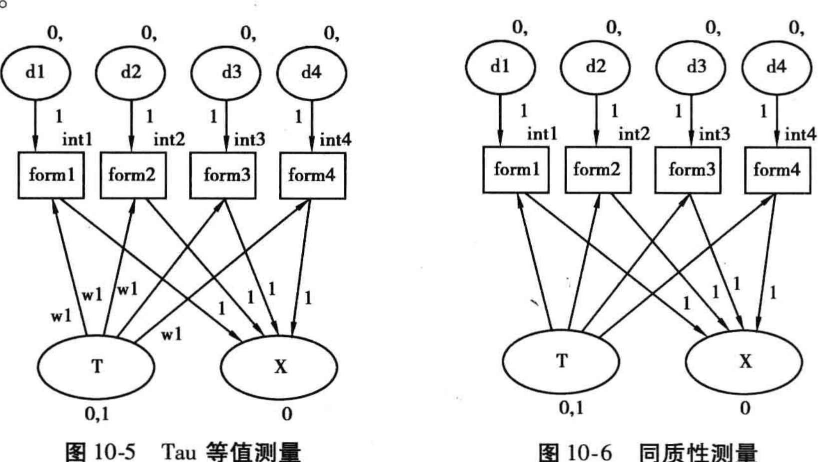


图 10-5　Tau 等值测量　　　　图 10-6　同质性测量

接着，研究者须点开模型管理窗口，确定三个测量模型的名称与相关参数的限制，参见图 10-8 到图 10-11。

三个测量模型中参数的等同限制，逐一列举如下。图 10-9 系平行测量模型，负荷量与误差变异量均设定为相等，注意这四个路径图中的相关路径系数，均设定为相等(v1 = v2 = v3 = v4；w1 = w2 = w3 = w4)。

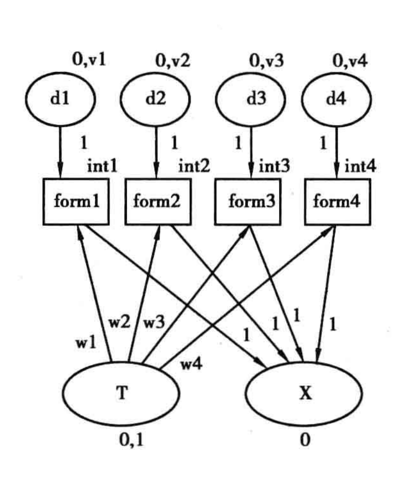


图 10-7　SEM 的信度分析模型

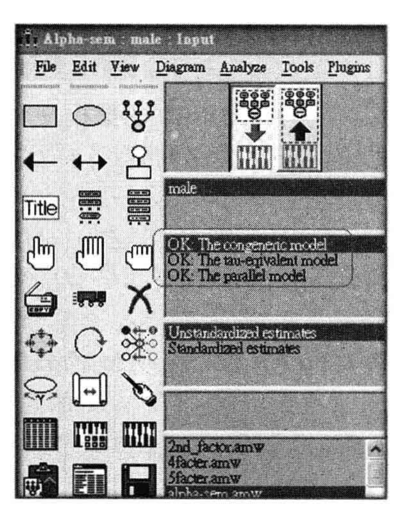


图 10-8　模型管理窗口的模型管理设定

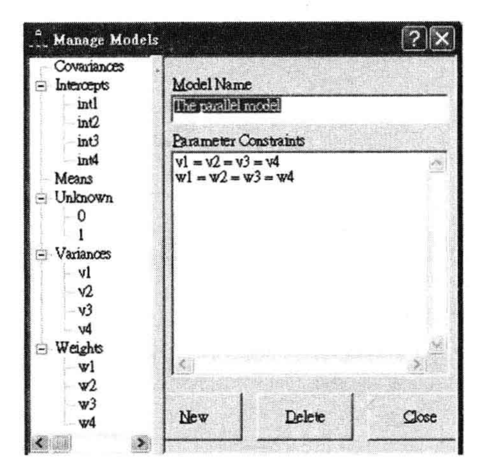



图 10-9　平行测量模型的参数等同限制

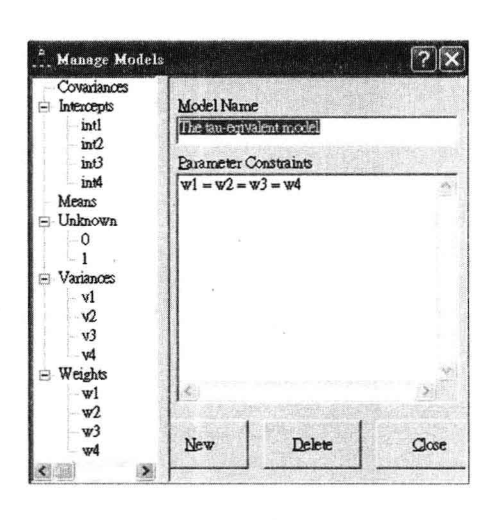


图 10-10　Tau-equivalent 模型的参数等同限制

图 10-10 系 Tau-equivalent 测量模型,负荷量设定为相等,注意这四个路径图中的相关路径系数,均设定为相等(w1 = w2 = w3 = w4)。

图 10-11 系同质性测量模型,本测量模型的参数均未等同限制,负荷量与误差变异量均开放估计,因此在[Parameter Constraints]的窗口中,未做任何设定。

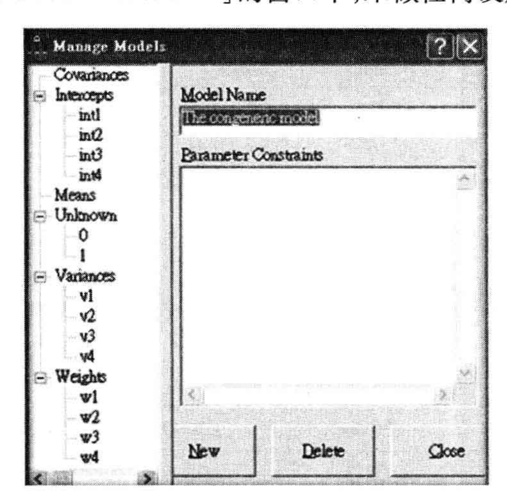


图 10-11　同质性测量模型的参数等同限制

由以上三个测量模型的设定可知,这些测量模型为嵌套设计(nested design),因此可以进行卡方差异的统计检验,以比较模型间的适配性。研究者须打开分析如图 10-12 的属性窗口,勾选[Standardized estimates]、[Squared multiple correlations]、[All implied moments]等相关参数,以便 Amos 能输出这些分析结果:题目信度指标(R^2)与隐含相关矩阵(Implied correlations)。

最后,利用原始数据文件(文件名 alpha-sem. sav)连接至 Amos 后,执行 Amos,得到表 10-2 中各测量模型的 Amos 适配度分析比较结果。

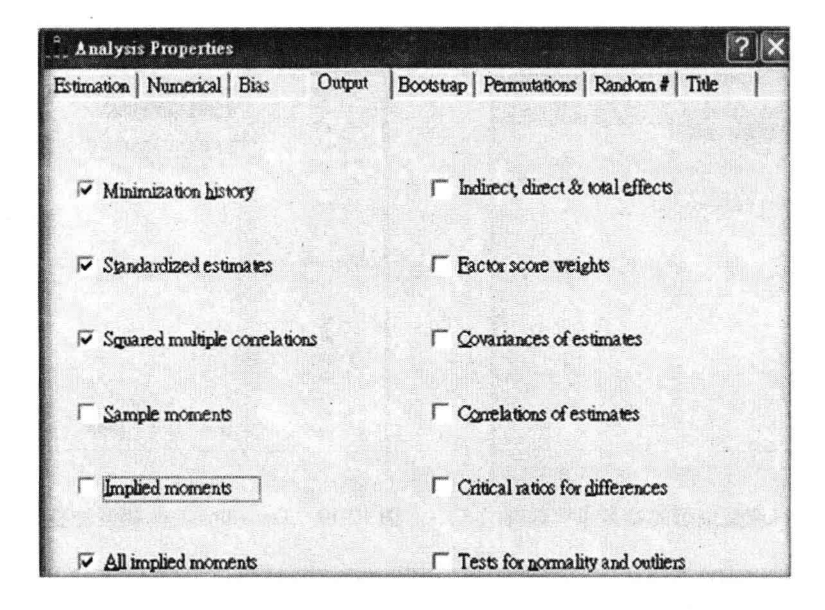


图 10-12　Amos 分析属性窗口中的基本设定

表 10-2　三个测量模型的适配度分析比较表

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
The congeneric model	12	2. 6156	2	. 2704	1. 3078
The tau-equivalent model	9	19. 5065	5	. 0015	3. 9013
The parallel model	6	34. 0253	8	. 0000	4. 2532
Saturated model	14	. 0000	0		
Independence model	8	270. 8184	6	. 0000	45. 1364

ECVI

Model	ECVI	LO 90	HI 90	MECVI
The congeneric model	. 0213	. 0209	. 0282	. 0214
The tau-equivalent model	. 0301	. 0221	. 0441	. 0301
The parallel model	. 0369	. 0254	. 0545	. 0369
Saturated model	. 0225	. 0225	. 0225	. 0225
Independence model	. 2300	. 1898	. 2762	. 2301

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
The congeneric model	. 0157	. 0000	. 0607	. 8727
The tau-equivalent model	. 0482	. 0269	. 0717	. 5057
The parallel model	. 0511	. 0341	. 0693	. 4253
Independence model	. 1881	. 1694	. 2076	. 0000

由表 10-2 中 P 值、RMSEA 、ECVI 等 SEM 适配度统计量观之,同质性测量模型(The congeneric model)的适配度均最佳(p = 0. 2704, RMSEA = 0. 0157, ECVI = 0. 0213)。再由表 10-3 的嵌套模型的比较结果知,三个测量模型间的差异均达 0. 05 的显著水平,再度确定了同质性测量模型的优越性。因此,要计算本测量的 Cronbach α 系数,在同质性测量模型下进行才是最佳选择。

接着,点选图 10-13 中 Amos Output 窗口的左侧[Estimates]下的[Matrices],接着点选[The congeneric model],打开同质性测量模型的隐含相关系数矩阵表。

表 10-3　三个测量模型间的嵌套模型的比较

Nested Model Comparisons

Assuming model The congeneric model to be correct:

Model	DF	CMIN	P	NFI Delta-1	IFI Delta-2	RFI rho-1	TLI rho2
The tau-equivalent model	3	16. 8910	. 0007	. 0624	. 0628	. 0575	. 0588
The parallel model	6	31. 4097	. 0000	. 1160	. 1168	. 0653	. 0667

Assuming model The tau-equivalent model to be correct:

Model	DF	CMIN	P	NFI Delta-1	IFI Delta-2	RFI rho-1	TLI rho2
The parallel model	3	14. 5187	. 0023	. 0536	. 0546	. 0078	. 0080

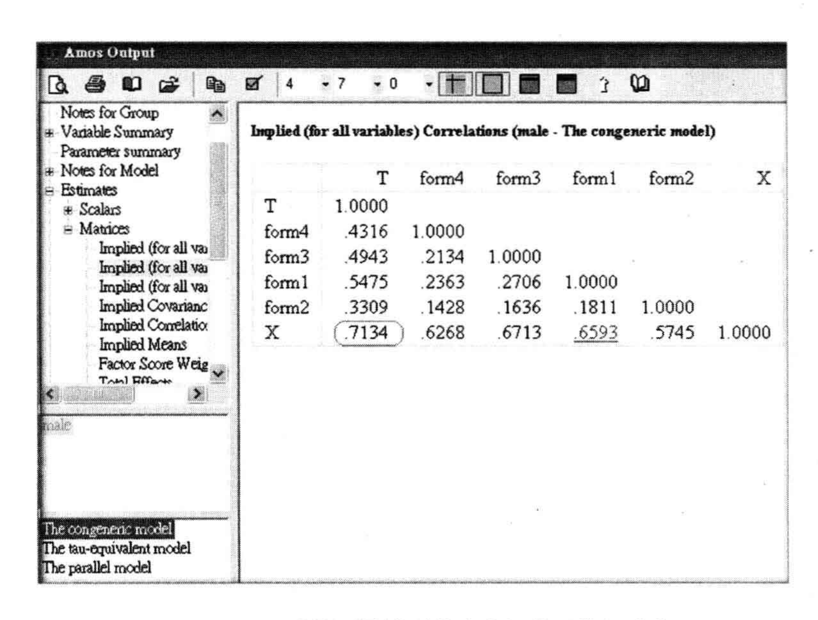
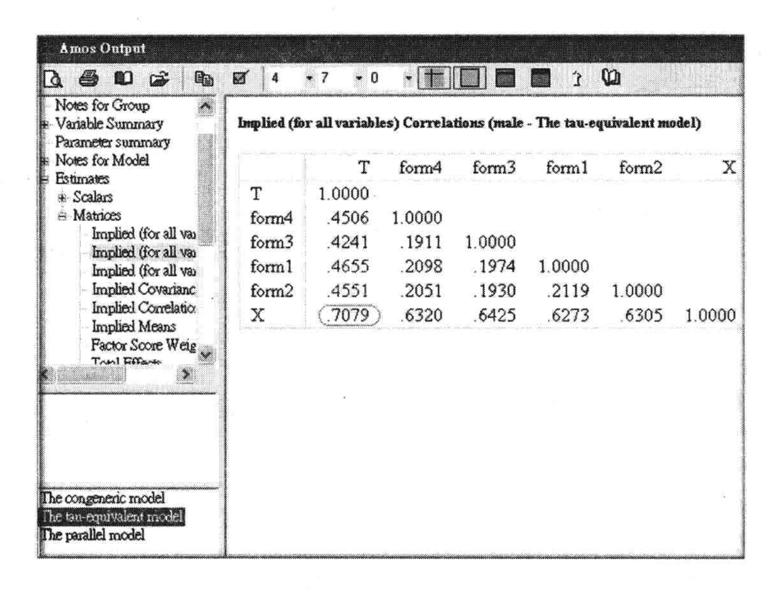


图 10-13　同质性测量模型的隐含相关系数矩阵表

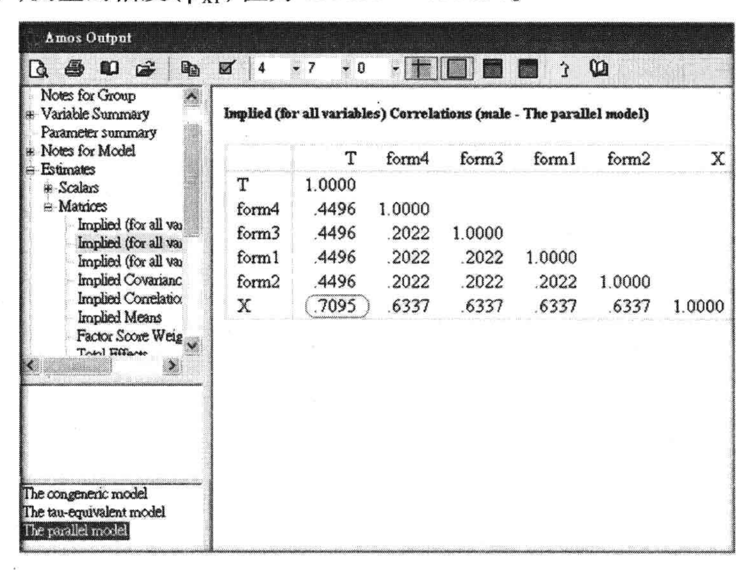
由上述的统计结果知,同质性测量模型的适配度最佳,因此本测量的信度(ρ_{XT}^2)的最佳估计值为 $0.7134^2 = 0.5089$,参见图 10-13 右侧窗口内容,否则会低估测量的信度(请比较表 10-1 中的 α 系数)。其次,打开[tau-equivalent]等同质性模型的隐含相关系数矩阵表,参见图 10-14 右侧窗口内容,其测量信度的估计值为 $0.7079^2 = 0.5011$。



	T	form4	form3	form1	form2	X
T	1.0000					
form4	.4506	1.0000				
form3	.4241	.1911	1.0000			
form1	.4655	.2098	.1974	1.0000		
form2	.4551	.2051	.1930	.2119	1.0000	
X	.7079	.6320	.6425	.6273	.6305	1.0000

图 10-14　Tau-equivalent 测量模型的隐含相关系数矩阵表

其次，再打开平行测量模型的隐含相关系数矩阵表，参见图 10-15 右侧窗口内容。在此测量模型下，测量的信度（ρ_{XT}^2）值为 $0.7095^2=0.5034$。



	T	form4	form3	form1	form2	X
T	1.0000					
form4	.4496	1.0000				
form3	.4496	.2022	1.0000			
form1	.4496	.2022	.2022	1.0000		
form2	.4496	.2022	.2022	.2022	1.0000	
X	.7095	.6337	.6337	.6337	.6337	1.0000

图 10-15　平行测量模型的隐含相关系数矩阵表

值得一提的是，在 SEM 的架构中，研究者还可建立建构信度（Construct reliability/Composite reliability），除了笔者（2006）过去所设计的程序外（参见附录四），Amos VB 程序设计亦是另一途径。以数常识测量中的一个分测量为例，直接利用 Amos 的 ProgramEditorCLI，在图 10-16 的窗口中撰写 Amos VB 程序（请参见 ConRel. vb），可获得该测量的建构信度为 0.6936。当然，使用此程序研究者需要具备一些基本的 Basic 语言运用能力。

按下图 10-16 中右下角的[Run]后，就会显示如图 10-17 中的建构信度。

另外，研究者如欲建立建构信度的置信区间，可利用 Amos Bayesian SEM 程序获取此

统计量。首先,以数常识测量中的一个分测量为例,建立理论模型如图 10-18,并设定所需参数的名称,以便在程序中引用。接着,连接数据文件(如文件名 NS94. sav)。

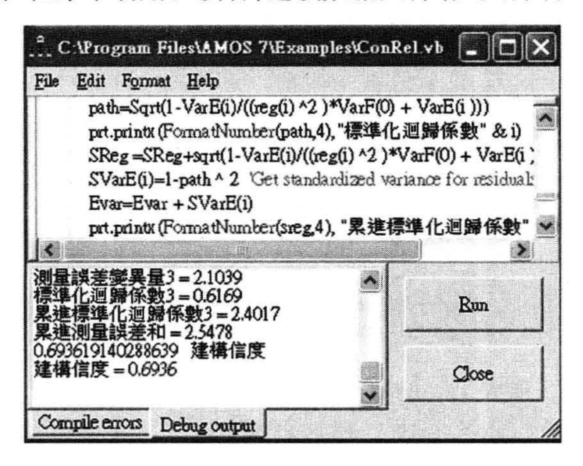


图 10-16　Amos Vb 的程序设计窗口



图 10-17　建构信度窗口

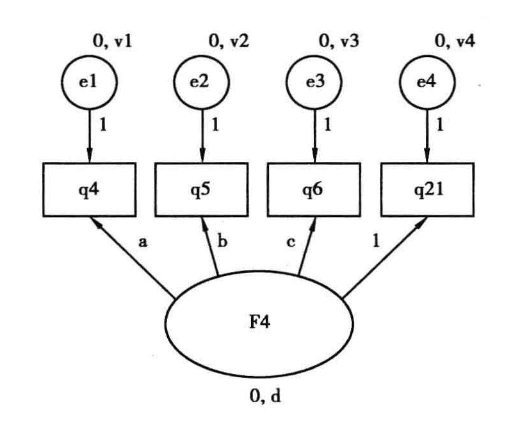


图 10-18　数常识分测量理论模型

其次,当研究者执行 Bayesian SEM 后,就会出现图 10-19 的 Bayesian SEM 的参数显示窗口,当笑脸出现后,即可按下自订参数估计钮,在图 10-20 窗口中撰写建构信度的程序,以计算建构信度的置信区间。图 10-19 的 Bayesian SEM 窗口中,(500 + 71501) * 4 表示 Amos 曾产生了 72001 * 4 个样本,但只保留了 71501 个样本, * 4 表示 Amos 已瘦身两次,因而只保留了 1/4,500 是 Amos 的内定 burn-in 样本数。

按下图 10-20 右下角[Run]之后,图 10-21 的自订参数估计值窗口,显示了 MCMC 的额外参数估计值:真分数方差(VarT)、误差分数方差(VarE)与建构信度(ConRel)。这三者的关系说明如下:设有一潜在因素 T,其观察变量为 X_i,未标准化回归系数为 b_i,则其观察变量 X_i的方差为 $Var(X_i) = b_i * Var(T) * b_i + var(\varepsilon_i)$。建构信度经由此关系可搭起与传统信度系数的桥梁。依照 Reuterberg 与 Gustafsson(1992)的定义:建构信度可利用未标准化的路径系数与潜在方差、测量误差方差求得。其公式定义如下:

Bayesian SEM

File　Edit　View　Analyze　Help

1.0016

(500+71,501) *4　1376　881

(500+71,500) *4

Group number 1

Regression weights	Mean	S.D.	C.S.	95% Lower bound	95% Upper bound	Name
q4<--F4	1.184	0.128	1.001	0.956	1.448	a
q5<--F4	1.323	0.146	1.001	1.070	1.652	b
q6<--F4	1.213	0.132	1.002	0.975	1.489	c

图 10-19　Bayesian SEM 的估计参数显示窗口

$$r_{TX} = \frac{Var(T)\left(\sum_{i}^{m} b_i\right)^2}{Var(T)\left(\sum_{i}^{m} b_i\right)^2 + \sum_{i}^{m} Var(\varepsilon_i)}$$

图 10-20 窗口内的程序设计，即依此公式撰写，程序中 VarT 代表潜在变量的方差，VarE 代表各指标的误差方差的和，Cov 代表指标路径系数平方和与潜在变量方差的乘积。

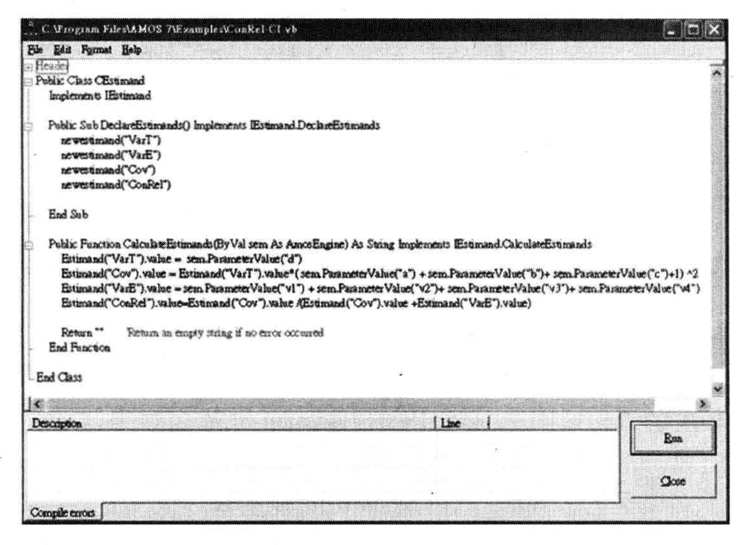


图 10-20　Amos VB 程序撰写窗口

在自订参数估计值图 10-21 窗口中，显示这个分测量的 0.95 置信区间为 0.6472 ~ 0.7280，71501 个 MCMC 样本建构信度的平均值为 0.6900，标准误为 0.0206，聚敛水平为 1.0009，前述建构信度的点估计值(0.6936)亦落在此置信区间内。

另外，在 SEM 的架构中，结构系数或相关系数会因测量误差而利用 Gulliksen (1950)的校正公式加以调整，研究者可善用此机制控制测量误差。例如，图 10-22 中，q1 ~ q3 为第一个分测量的分数，q4 ~ q6 为第二个分测量的分数，这两个分测量总分的

Custom Estimands

File Edit View Help

Numeric Estimands	Mean	S.D.	C.S.	95% Lower bound	95% Upper bound
VarT	0.9190	0.1616	1.0006	0.6274	1.2611
VarE	9.0488	0.3014	1.0007	8.4777	9.6634
Cov	20.2420	1.7141	1.0009	17.0281	23.7227
ConRel	0.6900	0.0206	1.0009	0.6472	0.7280

图 10-21 自订参数估计值显示窗口

Pearson 积差相关系数为 0.330(使用 NS55. sav 的原始数据文件),其 Cronbach 系数分别为0.373 与0.496。但是如果利用 Amos 建构下的模型设计,这两个分测量总分的相关系数(标准化路径系数)为 $0.767 (=0.330/(\sqrt{0.373}\sqrt{0.496}))$,与图 10-22 中 Factor 1 与 Factor 2 的相关系数相同。

不过此种测量误差对于路径系数的校正,无法考虑评分者差异或因测量时间等产生的测量误差,除非研究者将这些干扰因素纳入模型设计中。例如,图 10-23 中的评分者 SEM 路径模型即可控制评分者间的评定误差。

研究者如欲将研究人员(属于方法因子)纳入 SEM 的模型设计中,以检验测量方式的系统误差,可参考 Bollen(1993)探究自由民主的工具效度检验的做法,其路径图设计如图 10-24。图中“政治自由”与“民主规范”是潜在特质变量,其下各有四个观察指标,系来自不同研究者的观察指标。因此,图 10-24 下方的三个研究人员系观察指标的评分者:Sussman、Gastil 与 Banks,在此模型中被视为方法因子(method factor)。在此模型中,每一指标的变异量,可分割成三个部分:

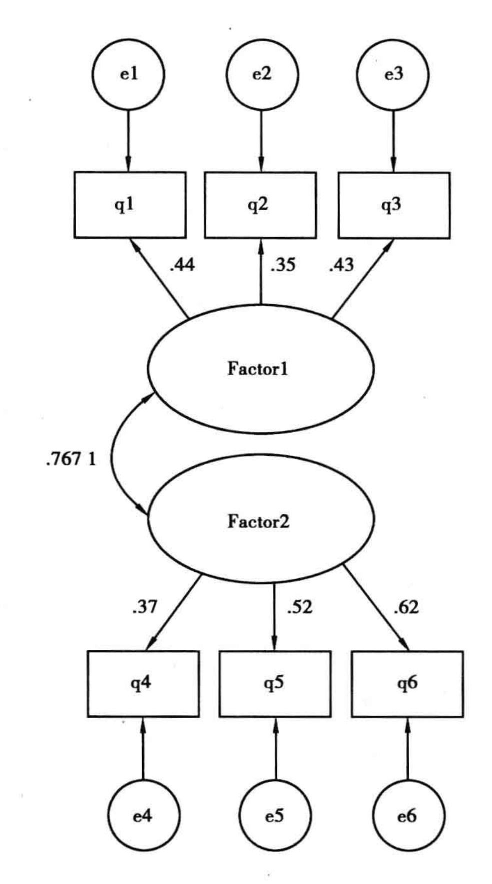


图 10-22 一般 CFA 设计(标准化路径系数)

1. 指标效度变异比:$\lambda_{ij}^2 Var(L_j)$;
2. 方法因子误差比:$\lambda_{ik}^2 Var(M_k)$;
3. 随机测量误差比:$Var(\delta_i)$。

其数学公式定义如下:$Var(X) = \lambda_{ij}{}^2 Var(L_j) + \lambda_{ik}{}^2 Var(M_k) + Var(\delta_i)$

例如,Var(群众抗争) = 1^2 · Var(政治自由) + 1^2 · Var(Banks) + E4

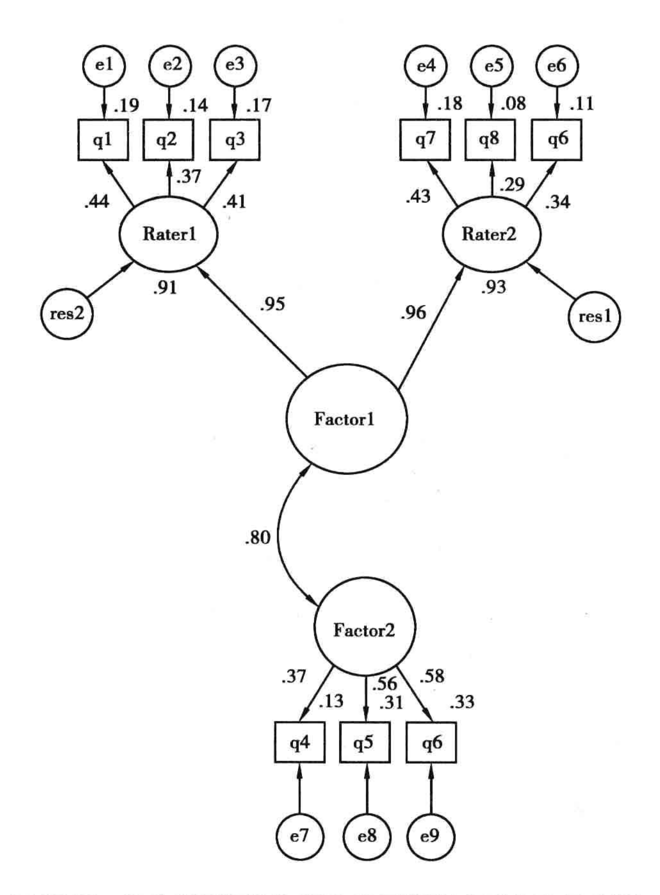


图 10-23　评分者评定误差 SEM 路径模型(标准化路径系数)

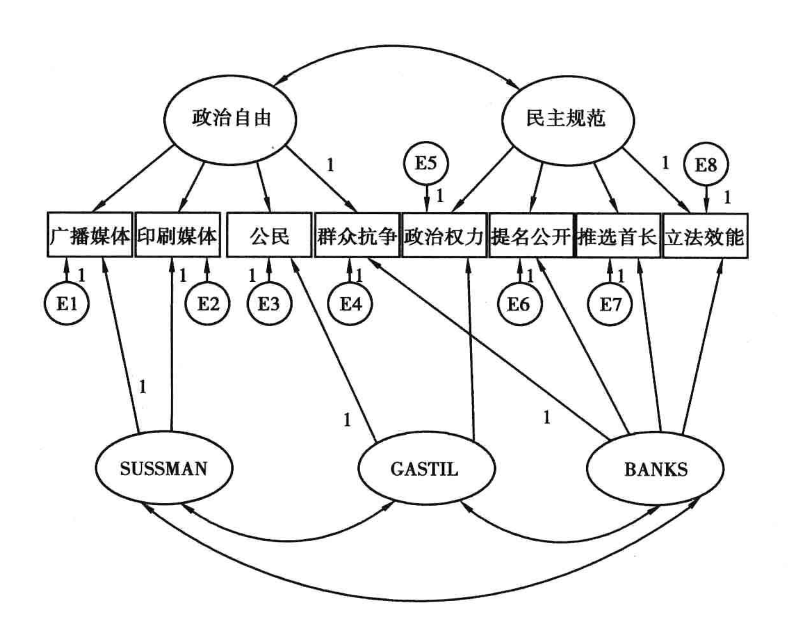


图 10-24　多重特质多重方法模型

利用上述的SEM模型,可以探究测量误差的不同来源(如评定人员特质、量尺类别、测量时间、数据取得方式等)及不同测量误差的大小,此种处理方式类似推论力理论(generalizability theory),可以将不同的变异源分开来估计。假如上述的评分者效果的确存在而您却将评分者因子拿掉的话,通常会使得观察指标间的随机测量误差产生高度相关。

过去有些研究者认为题目间具有高度相关就是同质性测量,而Cronbach α系数是单向度测量的指标,其实这种说法不完全正确。Hulin与Cudeck(2001 c & d)曾撰文就相关系数、Cronbach α与因素分析间的关系,做过详尽的比较与解析,值得量表编制者参考。他们认为:①无论测量题目多还是少,最好报告Cronbach α而非相关系数;②欲提高Cronbach α时,题目的品质比题数更重要;③Cronbach α并非单向度(unidimensionality)的指标,因为题目间具有高度相关并不一定是同质性题目,相关受到测量长度、题目间的相关、题目措辞的重复性与测量向度等因素的影响;④假如该测量是多向度的测量或研究者一味追求题目措辞的重复性,Cronbach α就可能偏高;⑤先进行因素分析再报告Cronbach α;⑥非单向度的量表有时亦很有用处,不要一味追求单向度的测量,例如组合性建构的量表。

五、效度(validity)的检验

传统常用的效度有:内容效度、效标关联效度、平均变异解释百分比与建构效度,较适合于单一维度建构且又是原因建构的测量。常用的统计方法有积差相关、探索式与验证性因素分析。更多的细节会在本章末第六节量表效度检验的内涵与方法中做说明。至于具体的应用实例分析请参考李茂能(2006)在另一Amos专书中讨论过有关量表信、效度检验的方法与实例,不在此另加赘述。

第五节　指标建构的过程

指标建构(index construction)乃指形成性指标的编制,编制时强调题目的独特性(特殊变异量)、降低多元共线性与多向度(Diamantopoulos & Winklhofer, 2001; Diamantopoulos & Siguaw, 2006)。由于指标建构的理论与实务到目前为止尚不够成熟,目前只能就各家的理论与相关的编制步骤,分阶段逐一整理后论述如下(Helm, 2005; Mackenzie, Podsakoff & Jarvis, 2005; Petter, Straub & Rai, 2007):

阶段一:数据搜集前

步骤1:内容面向的界定及观察指标的界定

内容面向的界定(Content specification)就是构念范畴与向度的厘清,彻底的文献探讨、利用焦点小组请教该领域专家或实务人员、实地观察都是可行之路。内容面向的界定亦需要广泛与周延,以确立内容效度。形成性指标的建构,其内容效度是指标建构效度最重要的内涵。这个核心步骤与前述量表发展的步骤并无二致。观察指标的界定(Indicator specification)与内容面向的界定密不可分,测量编制者须根据建构的每一内容面向,逐一据以编拟题目作为指标变量,漏掉某一面向即会改变潜在建构的组成。因此,指标的广泛性(comprehensive)与周延性(inclusive)仍须遵守。每一面向的观察指标的界定,依赖彻底的文献探讨、请教该领域专家与实地观察,广度重于深度,以免遗漏该建构中的重要层面而毁及内容效度。评估内容效度的方法有二:

1. 根据理论建构的层面或向度,评估该套指标是否完整地列述了该建构的领域或范畴。
2. 评估每一指标是否可作为理论建构的预测指标。研究者可邀请领域专家,将每一测量指标的卡片进行形成性建构或反映性建构的归类(Q-sorting or Expert Validation)。假如分类结果与事先界定的理论建构与指标间的关系具高度一致性,即显示具有内容效度(Straub, Boudreau & Gefen,2004)。

步骤 2:辨识建构的属性

研究者可以使用下列标准,判断自己的建构是否为形成性指标的建构(Diamantopoulos & Winklhofer, 2001; Jarvis et al. ,2003):

1. 利用指标预测建构吗?
2. 删除一个指标之后,会改变建构所测的内容吗?
3. 一个指标中的值改变,该建构下其他指标的值不会预期改变吗?
4. 建构中的指标有不同的前因与后果(different an tecedents and consequences)吗?

假如上述四个标准皆为真的话,那么您的建构肯定是形成性指标的建构;假如上述大部分的标准为真且又有理论依据的话,那么您的建构大概就是形成性建构。不过,如建构尚未很妥善定义的话,研究者可能一时无法回答前述的所有问题。此时,研究者的首要任务是再度检查界定该建构的领域或范畴,评估一下所有的指标在该特定建构下的适切性(Jarvis et al. , 2003)。

步骤 3:评估测量及结构模型

除了利用重测信度或路径系数评估题目的信度外,研究者须检查结构模型是否可辨识(Model Identification)(Jarvis et al. , 2003; MacCallum & Browne, 1993; MacKenzie et al. , 2005)。由于潜在建构及其误差项的测量量尺的未确定性,单独存在的形成性指标建构的 SEM 模型(如图 10-25-1),除非将构念层次的误差项设定为 0 或 1,在统计上常是无法辨识的(underidentified)。当使用协方差导向的 SEM 分析时,解决模型不可辨识的方法,常见的有五种(Petter, Straub & Rai, 2007):

1. 将一个(或一个以上)的结构路径系数或建构的误差项设定为 0。不过此种做法在理论上可能行不通(不可能没有测量误差),因此研究者只好改弦更张了,不能为了辨识而辨识。

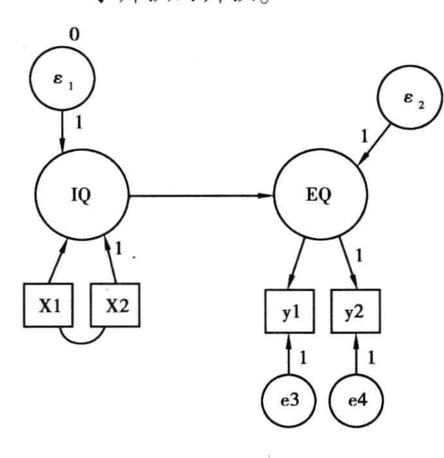


图 10-25-1 形成性指标单建构的辨识

2. 假如只有一条路径连接到一个反映性指标的建构,将形成性指标的建构加以解组(Decompose the formative construct),参见图 10-25-2 (b)。

3. 确保该形成性建构至少有两个以上结构路径系数,连接到反映性指标的建构,如图10-26(a)。虽然内部一致性(internal consistency)并不适用于形成性指标,但此时研究者仍须确认此一形成性建构是否具有外部一致性(external consistency),亦即此一形成性建构是否具有单一向度的特质而不具有点变量(point variable)的功能,如果不是则易导致解释性的困惑,造成此一形成性建构无法完全调节形成性指标与外衍建构间的关系(Howell,

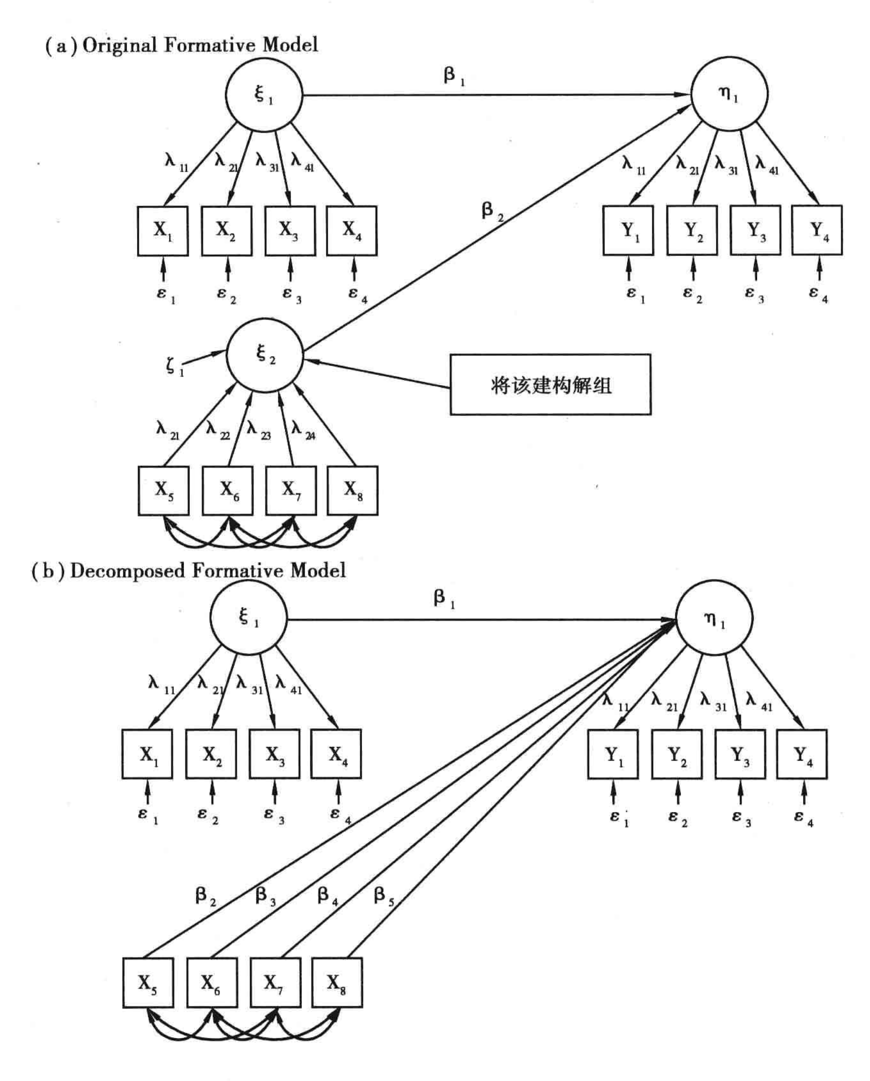


图 10-25-2　形成性指标建构的解组

资料来源:Petter,Straub 与 Rai(2007)

Breivik & Wilcox, 2007a & b)。

4. 假如概念上可行的话,在该形成性指标的建构内,包含假定没有测量误差的两个反映性指标:V4 与 V5(此即 MIMIC 模型:multiple indicators and multiple causes of a single latent trait),本质上这是反映性与形成性指标的混合模型(A hybrid model,参见图10-26(c)),Jarvis 等人(2003)认为这是最佳解决方案,主要理由为:(1)该建构可以就内衍建构或外衍建构,运用自如;(2)可将该建构纳入验证性因素分析模型中,评估其区别效度及测量的信度;(3)测量参数的估计值会更稳定。理想上,V4 与 V5 应选自最能代表形成性建构的标竿变量。

5. 在该形成性指标的建构内,包含一个反映性指标,且连接一个结构路径到一个反映性指标的建构(图 10-26(b))。假如图 10-26(b) 中 V4 反映性指标是一个标竿变量(marker variable)的话,研究者即可评估各形成性指标与标竿变量的间接效果,作为题目效度(item validity)的指标。

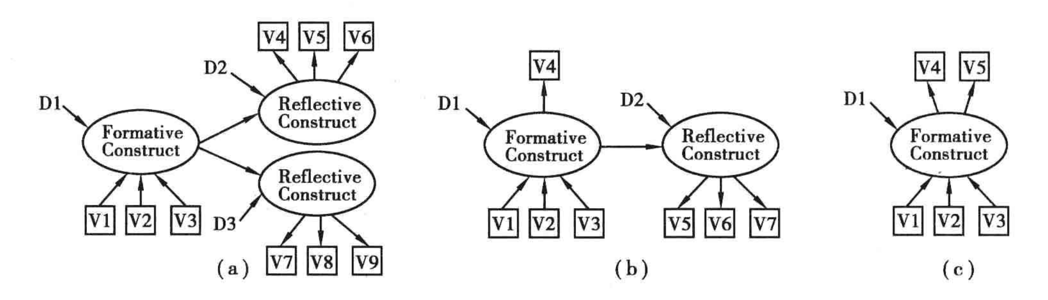


图 10-26　形成性指标测量模型的可辨识途径范例

资料来源:MacKenzie、Podsakoff 与 Jarvis(2005)

阶段二:数据搜集后

步骤 1:对形成性指标的信、效度与指标间共线性进行检验(indicator collinearity)

一个优良的形成性指标应具有以下几个特色:非常精简、显著的负荷量、能获得较佳的模型适配度和较高的解释变异量。评估形成性建构指标的效度时,为了保留测量指标的独特变异量,研究者须使用主成分分析而非共同因素分析,检查各指标的因素负荷量,以进行指标的筛选。研究者可以删除不显著的指标(Diamantopoulos & Winklhofer, 2001),但当危及到建构内涵的广度或模型的适配度时,亦可保留不显著的指标,以确保内容效度(Bollen & Lennox,1991)。

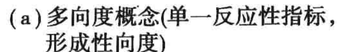
至于指标间共线性的检验,则是形成性指标必须通过的,因为这时指标是原因指标(causal indicators),通常为了精简与容易解释,并不希望指标间高度相关,否则会造成不必要的重复,而导致回归分析时的多元共线性问题。因此,指标间的多元共线性分析是指标建构的必经历程,研究者可以进行 VIF(variance inflation factor)的分析。(较严苛的标准:VIF <3.3(Diamantopoulos & Siguaw,2006);较宽松的标准可定为 VIF <10(Helm, 2005)。这与要求指标间须具有高度相关的量表发展是截然不同的。假如多元共线性的问题存在,常见的解决之道有:

1. 同时包含形成性指标与反映性指标;
2. 在内容效度不受影响的前提下,删除具有相关的指标;
3. 合并相关的指标作为单一的组合指标;
4. 转化成多向度的建构(multidimensional construct),参见图 10-27(a) 与(b)的做法。

前述"量表发展"着重各指标间的内部一致性,而"指标建构"则着重检验"重测信度"以评估各指标的稳定性,及观察指标与标竿变量或共通性题目(a global item)间的关系,以检验指标的外在效度(External validity)及建构的外部一致性(external consistency)。例如,各观察指标须与标竿变量或共通性题目间具有显著相关,即为指标的外在效度证据之一。所谓标竿变量或共通性题目系指与某一特定建构具有明确相关的变量,且其属性已在先前被确认过,能概括地反映待测建构的特质。举例来说,研究者如已新编制好一个四点量尺(几乎完全不会、有一点、经常、几乎完全会)的夜间睡眠质量的问卷。其问卷的五个题目内容如下:

请就以下几个问题评估一下在过去一个月以来,您夜间睡眠的困扰程度。

1. 上床后,您有无法入睡的困扰吗?
2. 夜间睡觉时,您经常会惊醒吗?



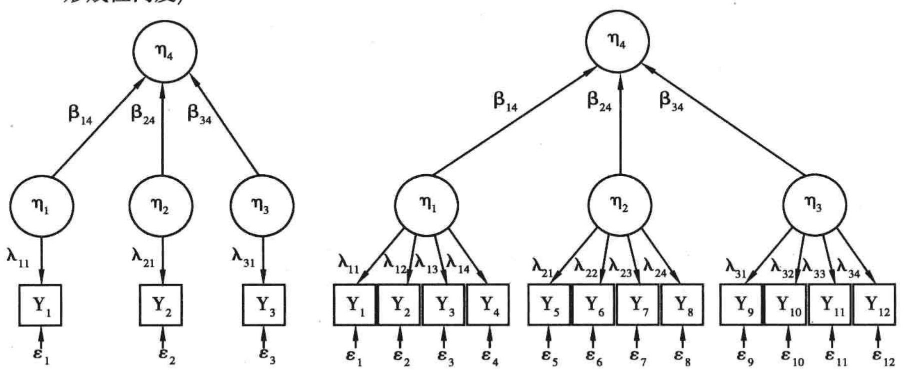


图 10-27　多向度的建构图解

资料来源:Petter,Straub 与 Rai (2007)

3. 夜间睡觉时,您感觉常躺在床上而睡不着吗?
4. 您常感觉太早醒来吗?
5. 您感觉夜间睡眠太少吗?

研究者此时就可编拟一共通性题目"整体而言,在过去一个月以来,您夜间睡眠的质量如何?",以检验指标的外在效度。而此一共通性题目的量尺,可采用以下的七点量尺:非常好、很好、尚好、不好不坏、有点差、很差、非常差。假如将此一共通性题目纳入因素分析中,它即为一标竿指标,可标定其他五个题目在因素空间上的位置,以确认新因素。

步骤 2:进行形成性指标建构的效度评估

形成性指标的建构效度评估,主要途径有三:

第一,可以通过协方差导向的 SEM(Covariance-based SEM)。例如,Amos,LISREL 或 Tetrad SEM: CTANEST1,分析形成性指标的建构模型在以下模型 A、B、C 间的适配性(Jarvis et al., 2003);亦即进行以下模型 A、B 与 C 间卡方差异性检验,以决定使用哪一模型。

1. 模型 A:外衍形成性指标的建构内的指标间具有共变关系,且这些指标又与其他外衍建构具有共变关系。
2. 模型 B:外衍形成性指标的建构内的指标间具有共变关系,但这些指标与其他外衍建构没有共变关系。
3. 模型 C:外衍形成性指标的建构内的指标间没有共变关系,且这些指标也与其他外衍建构没有共变关系。

此外,在 SEM 的分析架构下,亦须评估各指标的路径系数的大小与方向性,是否合理。

第二,可以通过成分导向的 SEM(Component-based SEM)分析软件(例如,PLS),分析形成性建构模型。

协方差导向的 SEM 分析与成分导向的 SEM 分析的关系,如同共同因素分析与主成分分析,两者最大不同在于前者追求模型的适配度,而后者追求解释变异量的极大化且不会有模型辨识性的问题(Chin, 1998; Petter, Straub & Rai,2007)。评估成分导向的 SEM 模型适配性,其步骤为(Chin, 1998):

1. 在研究模型中，界定形成性指标的建构（全部为形成性指标或混合式指标模型）；
2. 检查形成性指标的路径加权值或反映性指标的因素负荷量；
3. 评估结构模型中内衍变量（endogenous variables）的 R^2。

第三，研究者亦可以通过多重指标—多重原因模型（即 MIMIC 分析），进行建构网络效度分析。具体的实例，读者可以参阅 Diamantopoulos 与 Winklhofer（2001）、Diamantopoulos 与 Siguaw（2006）的论文。研究者通过此类 MIMIC 模型的适配度与路径系数，就可评估观察指标的贡献力与显著性，以挑选最适切的题目。另外，研究者亦可进行建构网络效度（Nomological validity）分析。此种建构网络效度是指待研究的建构在相关建构系统内，应有表现的程度（Cronbach & Meehl, 1955），这或可说是效标关联效度的变形，亦为外在效度的证据。因此，进行建构网络效度的分析，至少要有以下几个要素：至少要有两个建构、提出建构间的理论假说、能够加以测量的实征指标、指标变量与建构间的连接关系。研究者可根据理论，或相关研究假设，建立出构念与构念间及构念与外显行为间的系统网络关系，并绘出相关的路径图，以便通过 SEM 进行建构网络效度分析。其检验步骤为：①检验建构指标间的相关，以避免多元共线性问题；②评估形成性建构的测量模型是否可以辨识；③评估结构模型的适配度；④检查每一指标到组合变量的路径系数的强度与显著性。

综上所述，当建构是原因建构时，其指标必然为效果指标；当建构是效果建构时，其指标必然为原因指标。这两种测量模型，测量编制的方法与过程并不完全相同。因此，Bollen 与 Lennox（1991）认为传统的测量编制原则与理念应修正如下：

1. 同一建构的外显指标应具高度内部一致性的原则只适用于效果指标，不适用于原因指标（Cause indicators），各原因指标间不一定需要具有同质性（因为该建构可能为多维度的构念）。
2. 同一建构的外显指标间的相关愈高愈好的原则只适用于效果指标，不适用于原因指标，因为原因指标间具有高相关易导致多元共线性问题。
3. 单一维度建构中，信度相同的外显指标，本质上是可以替换的原则，不适用于原因指标。各原因指标本质上可能不同，删去任何一个原因指标都会破坏该建构的架构。
4. 不管是效果指标还是原因指标，建构内指标的相关不一定会高于建构间指标的相关（除非潜在变量间的相关为 0）。
5. 不管是效果指标还是原因指标，指标的线性组合都不等于潜在变量。

缘此，测量编制者在界定测量的建构时，首先须厘清该建构是原因建构（Causal construct），还是效果建构（Effect construct），亦即先根据所用理论厘清建构与指标间的因果关系，才能选择出最适切的观察指标。

第六节　量表效度检验的内涵与方法

量表的效度检验旨在确认研究中：①待测的特质是真实的、是可靠的；②待用的工具是否能真正测到所欲测量的特质，这是量表编制的核心任务。Cote 与 Buckley（1987）检验过去研究的 70 个 MTMM 矩阵数据，发现有 26.3% 的变异量来自于共同方法变异

(Common method variance),32% 的变异量来自于测量误差;Doty 与 Glick(1998)检视1980—1992 年六大社会科学期刊中的 MTMM 矩阵数据,发现有 32% 的变异量来自于共同方法变异,22% 的变异量来自于测量误差。由此观之,这些心理建构的测量工具仅能测到约 40% 的真正特质变异量,其余都是测量误差,不是系统性误差(如社会期许、方法/月晕效应、反应倾向),就是随机误差。为此,特辟此一专节深入论述,以利研究者编制出更有效的测量工具,进而提升量化研究的质量。不同形式的效度不应视为不同种类的效度,过去许多研究者(如 Gronlund, 1985; Crocker & Algina, 1986)将效度检验的策略,大致分为内容效度策略、效标关联效度策略、建构效度策略等三种,近来不少研究者(如 Straub, Boudreau & Gefen, 2004; Lewis, Templeton & Byrd, 2005)又增加了因素结构效度(factorial validity)与构念网络效度(nomological validity)。有些研究者甚至认为效度乃是单一的概念(a unitary concept),因而将内容效度、效标关联效度、因素结构效度(factorial validity)与构念网络效度均视为建构效度(Gronlund, 1985; Straub, Boudreau & Gefen,2004)。综合各家效度检验的策略(Churchill, 1979; Bagozzi, 1981; Grant & Davis, 1997; Malhotra & Grover, 1998; Straub, Boudreau & Gefen, 2004; Lewis, Templeton & Byrd, 2005; Garson, 2008),仍可细分为五类:内容效度、效标关联效度、区辨效度、聚敛效度与构念网络效度。研究者在选用这些效度证据时,亦应考虑到其在不同特性指标上的适用性。例如,内部一致性效度之聚敛证据就不适合于形成性指标上。兹依序说明这五类效度如下:

一、内容效度

内容效度(content validity)系指一个测量能均匀测到所欲测量内容领域的程度。内容效度包括三个重要的面向:①领域定义(domain definition);②领域代表性(domain representation);③领域关联性(content relevance)。后两个面向的评估通常视第一个面向的定义内容而定(Sireci, 1998)。领域定义是指内容领域的操作性界定,通常认知测量可通过双向细目表为之,情意测量则须通过文献分析为之。内容关联性系指测量中的每一题目应与建构范畴中的面向具有密切关联性,或测量内容与测量目的具有一致性;内容代表性系指量表能包含该建构所有面向的程度,或测量能够涵盖该内容领域定义的所有面向的程度,亦即建构的每一向度均编有代表性的题目,审查者可根据领域定义的内涵逐题评审。内容效度是其他效度的必要条件,它的检验应先于关联效度与建构效度,内容效度在成就测量上尤其重要。例如,研究者如欲利用成就测量来评估课程的学习效果,那么成就测量内涵与课程内容之间的适配性就是关键,否则测量的结果就可能偏颇而毫无意义可言。内容效度不管是经统计检验还是利用逻辑分析,一般均须先通过文献探讨、工作分析及该领域的理论或实务专家审定。Grant 与 Davis(1997)认为专家的"遴选及运用"关系着内容效度分析的效度,他们认为应该延聘富有经验的实务人员或理论专家至少二人以上,参与内容效度的分析。分析前亦应提供相关构念的理论架构或向度、告知该工具的测量模型:标准参照(强调构念的精熟度)或常模参照(强调题目的变异量)并清楚说明评审的三个要素:①题目的代表性(内容领域能充分反映构念的所有向度);②题目叙述的清晰性(题目的遣辞用字明确、读者易懂);③题目内容领域的完整性(全部的测量题目能反映所有的领域内涵)。

内容效度分析可为逻辑的或统计的分析,常用的题目内容效度的统计分析有:①评

分者间的一致性百分比指标；②Kappa 系数；③题目一目标一致性指标；④Aiken 效度指标；⑤Aiken 和谐系数；⑥Krippendorff's α；⑦Lawshe 内容效度比；⑧Q 分类法（Crocker, Miller & Franks, 1989；李茂能，1997）。兹依序简述如下：

（一）评分者间的一致性百分比指标

评分者间的一致性评估，通常研究者会邀请至少二或三位该领域的理论专家或实务人员，通过 E-mail 或问卷（含 Delphi 技术）就测量或量表题目内容的适当性、代表性、题目陈述的清晰性或选项的适切性，利用二分法加以评定，之后即可计算各题的评分者间一致性百分比；也可利用多点计分法计算各题的平均量尺值。一般来说，评分者间的一致性百分比应达 0.80 以上（Davis, 1992），各题的平均量尺值则至少应在 PR_{70} 以上。图 10-28 系看护工作负荷量表的内容效度分析的审查问卷，Grant 与 Davis（1997）用来说明内容领域专家如何就题目的代表性、题目叙述的清晰性与题目内容领域的相关性或完整性，进行内容效度审查。图中第一栏上半部系看护工作负荷的主、客观定义，下半部系看护工作负荷量表的样本题目，第二栏下半部系对题目是否具有代表性的评估，第三栏下半部系对题目归属向度的评估；底部的陈述则是对题目叙述清晰性的评估。

看护负荷题目	代表性	负荷向度
概念界定： 看护负荷：对于在家看护的要求与反应 看护负荷的客观定义：看护工作对于看护者家庭生活的改变或破坏程度 看护负荷的主观定义：看护者对于看护经验的态度或心理反应	1=该题对于看护负荷工作不具有代表性 2=该题需要大修对于看护负荷工作才能具有代表性 3=该题只需小修对于看护负荷工作就能具有代表性 4=该题对于看护负荷工作具有充分代表性	1=主 2=客 3=无法分类
1.提供我的家人个人看护(如洗照,喂食)	1　2　3　4 评论	1　2　3 评论
2.与我家人成员沟通	1　2　3　4 评论	1　2　3 评论
3.处理我家人的金钱与账单等问题	1　2　3　4 评论	1　2　3 评论
	清晰度：上述这些题目书写清晰,各题皆具独特性，且能为看护者所理解吗？ ——是的,以下这些题目文意清晰 ——不是,以下这些题目文意不清晰 建议：	

图 10-28　看护工作负荷量表内容效度的审查问卷

资料来源：Grant 与 Davis（1997）

在本例中,各题的代表性的平均量尺值应在3.0以上,评分者对于题目预定向度的归属的一致性百分比应达0.80以上,否则该题目应考虑加以修正或删除。

图10-29是另一检验题目与内容领域相关性的查账员(auditor)检核表,评鉴时可邀请一些相关领域的专家评估各题目与内容领域的相关性。请评分者利用以下的十点量尺,针对各测量题在四个内容领域的相关性予以评分。

1 2 3 4 5 6 7 8 9 10

完全无关 高度相关

专业责任 内部管控 证据与程序 报告

1.查账员责任的范畴与特性

一般会在下列项目中明列:

(a)资方信件

(b)查账报告范围

(c)雇用信件

(d)查账报告简介

图10-29 查账员考试题目与内容领域的相关性检核表

资料来源:Sireci(1998)

本例的内容领域的相关性评估采十点量尺,相关领域的专家利用此量尺评估各题目在四个内容领域上的相关性程度。各题目与预定所归属的内容领域相关性的平均量尺值应在70.0以上,否则该题目应考虑加以修正或删除。

(二)Kappa 系数法

Kappa系数法适合于归类一致性的分析,它是评分者实际评定一致性百分比与理论上评定一致性百分比的比值。SPSS菜单[Analyze]下的SPSS CrossTab子程序可以计算Kappa系数,但仅适用于两位评分员。图10-30中的数据文件,系SPSS CrossTab建档格式,研究者须在SPSS Data Editor中建立三个变量:rater1,rater2与freq,其中freq变量系用来做数据加权。

由图10-30中的次数分布可看出,整体而言,这两位评分者在这20道题目上,认为有必要纳入量表的比值比较高($p = 15/20 = 0.75$),但亦有一题落入"没有必要"可以删除。注意,假如研究者使用原始数据而非整理过的数据,就不需要在CROSSTAB分析前进行加权。此时,原始数据就会有20笔数据。

表10-4 SPSS CrossTab 报表

	Value	Asymp. Std Error[a]	Appox. T[b]	Approx. Sig.
Measure of Agreement Kappa	.701	.200	3.807	.000
N of Valid Cases	20			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

由表10-4知,排除掉偶然性后,评分者间信度为0.701($p = 0.000$)。很可惜,目前SPSS还无法分析两位以上的评分者间信度。

KAPPA.SAV - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

13 : rater1

	rater1	rater2	freq	var
1	必要	必要	15	
2	必要	有用但非必要	1	
3	必要	沒有必要	0	
4	有用但非必要	必要	1	
5	有用但非必要	有用但非必要	2	
6	有用但非必要	沒有必要	0	
7	沒有必要	必要	0	
8	沒有必要	有用但非必要	0	
9	沒有必要	沒有必要	1	
10				

图 10-30 SPSS CrossTab 数据输入格式范例

（三）题目—目标一致性指标

题目—目标一致性指标(item-objective congruence index),由 Hambleton(1980)创用,特别适用于效标参照测量(Crocker, Miller & Franks,1989)。通常使用 +1,0, -1 三点量尺,评估每一题目能测到的各个教学目标的程度,其中, +1 表示该题目确定能测到该教学目标,0 表示该题目不确定能测到该教学目标,而 -1 表示该题目确定不能测到该教学目标。题目 i 与教学目标 k 间的一致性指标的计算公式为:

$$I_{ik} = \frac{N}{2N-2}(\mu_{ik} - \mu_{i..})$$

式中,N 表示教学目标个数;$\mu_{i.k}$表示评分者对于题目 i 与某教学目标 k 间一致性评定的平均值;$\mu_{i..}$表示评分者对于题目 i 与所有教学目标间一致性评定的平均值。注意本指标假设每一题目只测量单一目标。以表 10-5 中第一题与第二个教学目标间的一致性评定为例,$\mu_{1.2} = 6/6 = 1$,表示所有评分者均一致认为第一题是在测量第二个教学目标,而:

$$\mu_{1..} = \frac{18-9}{30} = \frac{3}{10} = 0.30$$

代表评分者对于题目 1 与所有教学目标间的一致性平均值。因此,第一题与第二个教学目标的一致性指标等于:

$$I_{1,2} = \frac{5}{8}(1-0.30) = 0.44$$

该指标值似乎偏低,因为该题被认为亦在测量其他目标(如第三或第四个教学目标)。

题目—目标一致性指标的值介于 ±1 之间, +1 表示所有的评分者均同意该题目 i 仅在测某一教学目标 k,假如某一题目被评定在测一个以上的教学目标,其题目—目标一致性指标的值就会小于 1。一般来说,题目—目标一致性指标的值大于 0.70 以上,才可视为一致性佳(Sireci, 1998)。

表 10-5　题目—目标一致性指标的数据文件

Item	Judge	Objective 1	2	3	4	5
1	J-1	–	+	+	+	0
	J-2	–	+	+	+	0
	J-3	–	+	+	+	–
	J-4	–	+	+	+	0
	J-5	–	+	+	+	–
	J-6	–	+	+	+	–
2	J-1	+	–	–	–	–
	J-2	+	–	–	–	–
	J-3	+	–	–	–	–
	J-4	+	–	–	0	–
	J-5	+	–	–	–	–
	J-6	+	–	–	–	0
3	J-1	–	–	0	0	0
	J-2	–	0	–	0	0
	J-3	0	–	–	0	+
	J-4	–	–	0	–	0
	J-5	–	0	–	–	0
	J-6	0	–	–	0	–

(四)Aiken 效度指标

Aiken(1980)提出一个题目适当性或关联性的指标,亦可作为内容效度指标,其计算公式如下:

$$V = \frac{\sum_{i=1}^{c-1} i n_i}{N(c-1)}$$

式中,N 表示评分者人数;c 表示评定量尺的类别数;i 表示各类别的加权量(最低类别的加权量为 0,次类别给予加权量 1,以此类推,最高类别的加权量为 c-1);n_i 表示将某一题目放入某一类别 i 的评分者人数。假设研究者请来六位评分者,使用三点量尺:0 表示不相关,1 表示不确定,2 表示确定有相关,在 0, 1, 2 三个类别上得到 1, 3, 2 的次数分布,其 Aiken 效度指标为:

$$V = \frac{1(3) + 2(2)}{6(2)} = 0.58$$

V 值的随机概率值计算公式为：

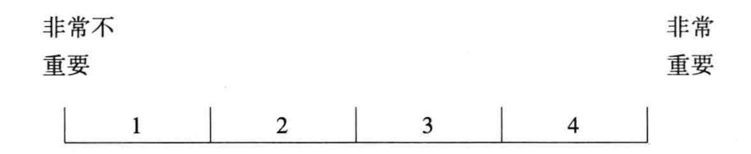
$$p = \frac{\frac{N!}{c^N}}{n_0!n_1!\cdots n_{c-1}!}$$

$$p = \frac{\frac{6!}{3^6}}{1!3!2!} = 0.01$$

假如评分者只是随机评定的话，获得 0.58 的效度值的概率为 0.01。

（五）Aiken 和谐系数

笔者（1997）所介绍的 Aiken 和谐系数，亦可用来评估及检验每一题目的评分者一致性，并作为选题的依据，文中亦附有 SAS 程序可资运用。使用此方法时，建议邀请数位专家评分者，针对问题“本题目测到待测建构的重要性”，根据以下评定量尺，对每一问题做重要性同意程度的判断：



Aiken's H 系数亦适用于评分者一致性分析，而且又有显著性检验方法，可以检验其结果的发生概率。它可计算 N 位评分者对题目 j 的信度指标（interrater reliability）。Aiken 氏评分者间信度（H 系数）的计算公式为：

$$H_j = 1 - \frac{4S_{.j}}{(c-1)(n^2-k)}$$

$$S_{.j} = \sum_{i'=1}^{n-1} \sum_{i'+1}^{n} | x_{ij} - r_{r'j} |$$

上式中，$j = 1, \cdots, m$ 个题目；r_{ij}为评分者 i 在题目 j 上的评定结果；n 表示评分者人数；c 为量尺类别数；k 为虚拟变量。当 n 为偶数时，$k = 0$；当 n 为奇数时，$k = 1$。从公式一得知，每一评定问题均可求得一 Aiken H 系数，参见表 10-6 的最后一栏 H 上的 Aiken H 系数。该系数介于 0～1，当评分者均给予相同评分时，该题评分者信度为 1，评分者间的评分差距越大时，该题评分者信度便会越趋近于 0。所求得的信度还可通过统计方法加以检验，检验公式如下：

$$z = \frac{\overline{H} - \mu_H}{\frac{\delta_H}{\sqrt{m}}}$$

式中，$\overline{H}$表示整体 Aiken H 平均系数。

表 10-6 是 7 位评分者对于 15 题有关计算机软件的评定结果，5，11，14 题的评分者间信度似乎偏低，需考虑加以修正或删除。表 10-6 底部，列出全量表 Aiken H 的平均系数，与 Aiken H 系数的 Z 检验值。一般来说，Z 检验值都希望能大于 1.9645。

表 10-6　第一套计算机软件评定结果的信度分析

题目	R1	R2	R3	R4	R5	R6	R7	x（总分）	H
1	3	4	4	4	4	5	5	29	0.667
2	4	4	4	4	4	5	5	30	0.792
3	4	4	4	4	4	5	5	30	0.792
4	3	3	3	3	4	5	5	26	0.542
5	2	3	2	3	4	4	5	23	0.417
6	4	3	4	3	4	5	5	28	0.583
7	1	1	1	1	2	1	1	8	0.875
8	1	1	1	1	2	1	1	8	0.875
9	1	1	1	1	3	1	1	9	0.750
10	1	1	1	1	3	1	1	9	0.750
11	2	2	2	2	4	4	1	17	0.458
12	3	2	3	3	4	4	5	24	0.500
13	2	2	2	2	4	2	1	15	0.625
14	2	2	2	2	4	4	5	21	0.375
15	4	3	3	4	5	4	5	28	0.583
$r_{i.x}$	.924	.924	.933	.958	.775	.938	.865		

$\overline{H}=0.64$, $\mu_H=0.3$, $\delta_H=0.15$, $Z_H=8.59$.

（六）Krippendorff' s alpha

Krippendorff' s alpha(α) 系数由 Krippendorff 所创用(Hayes & Krippendorff,2007),适合于各种测量层次的评分者间信度的评估,本系数不受评分者人数、评定类别数或量尺分数多寡的限制,还可计算置信区间,可说是多才多艺。在内容效度的评估上,Krippendorff' s α 可谓后起之秀。其计算公式:$\alpha=1-\frac{D_o}{D_e}$,式中 D_o 表不一致的观察值,D_e 表不一致的期望值。由于计算的过程较繁复,研究者可以下载 Hayes 与 Krippendorff (2007)所研发的 SPSS 宏(网址:http://www. comm. ohio-state. edu/ahayes/macros),研究者执行此宏之后,仍以表 10-6 数据为例(建文件格式亦同),在 SPSS 的 Syntax 窗口中输入以下程序语法:

```
KALPHA JUDGES = R1 R2 R3 R4 R5 R6 R7/LEVEL = 2/DETAIL = 1/BOOT = 100.
```

此行指令中,KALPHA 系宏的名称,JUDGES 之后是需要界定评分者的变量名称,LEVEL 后是需要界定的测量层次:1 表示名义变量、2 表示次序变量、3 表示等距变量、4 表示比率变量,DETAIL = 1 是要求输出 coincidence 与 delta 矩阵,BOOT = 1 000 是要求 Bootstrap 样本数,以便计算 Krippendorff' s α 的置信区间。接着,按下 SPSS 语法执行按钮,即可获得表 10-7 的统计结果。

表 10-7 SPSS Krippendorff's α 报表

Krippendorff's Alpha Reliability Estimate

	Aplha	LL95% CI	UL95% CI	Units	Observrs	Pairs
Ordinal	.6392	.5805	.6963	15.0000	7.0000	315.0000

Number of bootstrap samples:
1000
Judges used in these computations:

R1	R2	R3	R4	R5	R6	R7

由表 10-7 知,Krippendorff's α 为 0.6392,其 0.95 置信区间为 0.5805～0.6963,与 Aiken H 平均系数(0.64)相当接近。

(七)Lawshe 内容效度比

对于内容效度的评估,Lawshe(1975)建议邀请数位专家评分者,针对以下问题,逐一做判断:

本题目测到待测的知识或技巧的必要性。

1. 必要(Essential);
2. 有用但不必要(Useful but not essential);
3. 没有必要(Not necessary)。

假如有一半以上的评分者认为该题很必要或有用(如采较宽松标准),该题即被视为具有内容效度,能反映该领域的范畴。Lawshe(1975)的内容效度比(Content validity ratio,简称 CVR)公式:

$$CVR = (n_e - \frac{N}{2})/(\frac{N}{2})$$

式中,n_e = 认为该题很重要的评分者人数,N = 全部评分者人数。为了确保评分者间一致性而将偶然性纳入考虑,请参照表 10-8 的评分者人数及 CVR 的随机最低值(α = 0.05,进行观察指标的解释。

表 10-8 评分者人数及 CVR 的最低标准

评分者	CVR 最低值
5	.99
6	.99
7	.99
8	.78
9	.75
10	.62
11	.59
12	.56
13	.54
14	.51
15	.49
20	.42
25	.37
30	.33
35	.31
40	.29

计算出来的 CVR(介于 -0.99 ~ +0.99),会出现以下几种结果:

负值 CVR:少于一半的评分者认为该题有必要或有用;
零:刚好有一半的评分者认为该题有必要或有用;
正值 CVR:多于一半的评分者认为该题有必要或有用。

例如,六位评审专家中仅有三位认为某一题目是必要的,则该题的 CVR = (3 - 6/2)/(6/2) = 0,亦即刚好有一半的评分者认为该题有必要,如有五位认为必要,则该题的 CVR = (5 - 6/2)/(6/2) = 0.67,亦即多于一半的评分者认为该题有必要。

由表 10-8 的最低指标的随机值来看,当评分者人数小于 8 人时,只有当所有评分者均肯定该题的重要性时,该题才能存活下来。评分者人数越多,越有可能降低 CVR 的偶然性。因此一般来说,评分者人数介于 8 ~ 12 人,可能是比较切合实际的选择。

※实例应用

Lewis 等人(1995)利用 Lawshe 内容效度比,评估他们的测量工具(information resource management, IRM)的题目内容。在该研究中,评分者逐题评估各题的必要性或有用性,研究者计算各题的 CVR,以作为题目去留的指标。

(八)Q 分类法

Q 分类(Q-sort)法,由英国学者 Stephenson(1953)所创用,旨在针对个体的意见或知觉进行分类。通常分为三个阶段:Q 样本题目的建立、Q 分类的实施、因素分析 Q 分类。Q 分类乃利用个体的知觉与意见的辨识与分类,检视个体间的冲突、摩擦或共识的所在,因素分析则在于根据相似想法或意见进行分组。目前已有免费的统计软件 PQMethod 可资运用,深具效力与便利。下载网址:http://www.lrz-muenchen.de/~schmolck/qmethod/downpqx.htm。

不过,一般在使用 Q 分类验证内容效度时,通常仅限于简单计算分类的一致性。因此,Q 分类的复杂统计细节在此不赘述。

※实例应用

Segars 与 Grover(1998)使用 Q 分类检视 SISP(strategic information systems planning)的内容建构。经过文献探讨,他们发现 SISP 含有四个向度:(1)alignment;(2)analysis;(3)cooperation;(4)improvement capabilities ,以及 28 个相关的计划目标。他们将 SISP 的 28 个计划目标分别随机地写在纸条上,请专家将这些目标分类到四个相关的向度中。整体的分类正确百分比为 82%,研究者决定各题的分类正确百分比须超过 90%,才会被保留下来。结果共有 23 个计划目标最后被用做四个向度的测量指标。

二、效标关联效度

效标关联效度包含同时效度(concurrent validity)、预测效度(predictive validity)与处方效度(prescriptive validity)。同时效度、预测效度与处方效度均为构念网络效度(Nomological validity)的一部分。同时效度是指单一变量与同时存在的外在效标间的关系程度;而预测效度是指单一变量与其前因或后果变量间的关系程度;俗称的处方效度是指该量表诊断身心功能的程度。效标关联效度的统计分析实例,请参见聚敛效度与构念网络效度相关章节。

三、区辨效度

区辨效度(discriminant validity)指不同构念内的指标间不应具有高相关,以及不同构念或建构间相关要低。合乎此要求,才不会让人觉得这些不同构念的指标在测相同的构念或这些不同的构念的重叠性过高。Ping(2005)提议通过以下方法评估区辨效度:(1)相关分析(如来自两个不同构念的指标间相关大于 0.85,即为删题的可能对象,不同构念间相关应小于 0.70);(2)因素分析;(3)AVE 方法(average variance extracted);(4)SEM 方法;(5)MTMM 方法(multitrait-multimethod)。兹依序说明如下:

(一)相关法

以一般经验法则来说,两个构念间的相关系数如小于 0.70,即视为具有区辨性(Bagozzi ,Yi,1988);Anderson 与 Gerbing(1988)则主张利用 $\Phi \pm 2SE$ 的置信区间来判断两个建构的相关系数的 CI 是否包含 1,假如未包含 1 即表示这两个建构具区辨效度,亦即两个建构的相关系数显著小于 1。相关实例,请参见表 10-16 的正文说明。

(二)AVE 方法

虽然 AVE(average variance extracted)常作为聚敛效度的指标,但 Fornell 与 Larcker(1981)提议检查各构念内的 AVE 是否大于各构念间的决定系数,可凭此检验区辨效度。假如构念间的决定系数(r^2)(i. e. ,共同分享变异量)小于个别构念的 AVE(扣除独特误差的变异量),即显示出个别构念所抽取的变异量大于它们所共同分享的变异量(r^2)。换言之,假如个别构念内部的相关大于个别构念间的决定系数(较严苛)或相关系数,构念间就具有区辨性。参看表 10-9 中 Goo, Kim 与 Cho(2006)的应用实例:Structure of Service Level Agreements(SLA) in IT Outsourcing: The Construct and Its Measurement。表 10-9 中,对角线之外的元素就是各个构念间的相关系数。

表 10-9 个别构念内部的相关与个别构念间的相关矩阵

	SLO	POP	SLC	ACP	FDMP	IP	FP	CP	MC	CAC	EP
SLO	0.83										
POP	0.52	0.89									
SLC	0.60	0.60	0.87								
ACP	0.35	0.55	0.51	0.90							
FDMP	0.39	0.61	0.61	0.65	0.90						
IP	0.37	0.41	0.45	0.66	0.67	0.86					
FP	0.44	0.48	0.54	0.56	0.69	0.61	0.91				
CP	0.48	0.58	0.64	0.47	0.56	0.46	0.63	0.90			
MC	0.53	0.57	0.70	0.40	0.62	0.50	0.64	0.69	0.91		
CAC	0.42	0.38	0.45	0.50	0.55	0.56	0.64	0.54	0.59	0.85	
EP	0.25	0.43	0.44	0.21	0.34	0.28	0.29	0.43	0.31	0.32	0.87

注:Bolded diagonal elements are square root of average variance extracted (AVE). These values should exceed the inter-construct correlations (off-diagonal elements) for adequate duscrumunant valldity.

SLO = Service Level Objectives; POP = Process Ownership Plan; SLC = Service Level Content; ACP = Anticipated Change Plan; FDMP = Future Demand Mgmt Plan; IP = Innovation Plan; FP = Feedback Plan; CP = Communication Plan; MC = Measurement Charter. CAC = Conflict Arbitration Charter; EP = Enforcement Plan.

表 10-9 中,对角线上数字是各个构念内的 AVE,AVE 的定义(Fornell & Larker,1981)如下:

$$AVE = \frac{(\sum \lambda_i^2)\mathrm{Var}(X)}{(\sum \lambda_i^2)\mathrm{Var}(X) + \sum \mathrm{Var}(e_i)}$$

式中,λ_i指未标准化路径系数,Var(X) 指方差,Var(e_i) 指测量指标 i 的误差方差。根据表 10-9 ,显然各构念内的 AVE(对角线系数)均大于各构念间的相关系数,因此构念间具有区辨效度。

(三) 因素分析

通过探索式因素分析,检查各构念内的指标是否出现跨因子的现象,即不仅每一测量指标要显著地落在单一因素上,通常希望该因素负荷量大于 0.60,而且在其他因素上每一测量指标的因素负荷量应小于 0.30。

(四) SEM 途径

设有两个构念 IQ 与 ACH,其相关的路径如图 10-31,如欲利用 Amos 来检验这两个建构或构念间的区辨性,可在模型的管理窗口中建立一个开放模型(The Unconstrained Model):方差设定为 1(V1 = V2 = 1),但协方差未限制。再建立一个限制模型(The Constrained Model):构念间的相关设定为 1(协方差与方差均限制为 1,亦即 V1 = V2 = CO = 1),相关的参数限制参见图 10-31 内部小窗口的说明。假如这两个模型的卡方差异值未达显著差异,即表示这两个构念不具备区辨效度,即无法拒绝 H_0: Φ = 1 的假设。此种区辨效度的 SEM 分析,研究者一次仅能进行一对构念的检验。

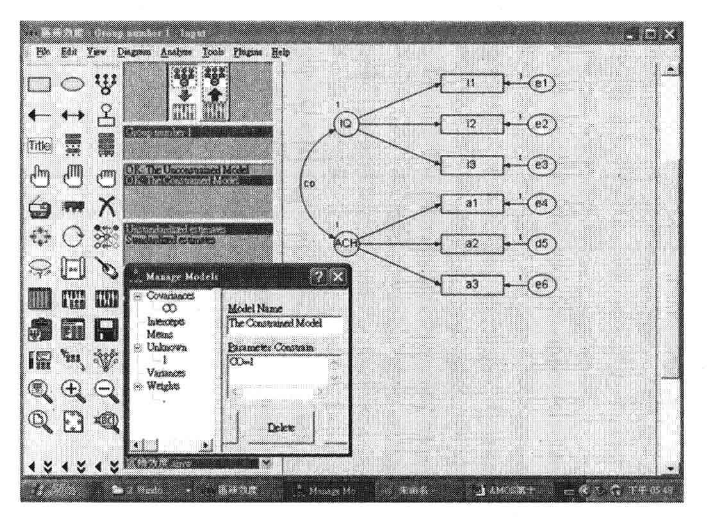


图 10-31　区辨效度的 SEM 分析路径模型

图 10-32 中 IQ 与 ACH 间的标准化路径系数(即相关系数)值与表 10-10 的估计值甚为接近,假如图 10-32 中的小数位数增到四位,两个估计值就会相等。

图 10-33 中 IQ 与 ACH 的路径系数设定为 1,其余参数是受限模型下的分析结果。

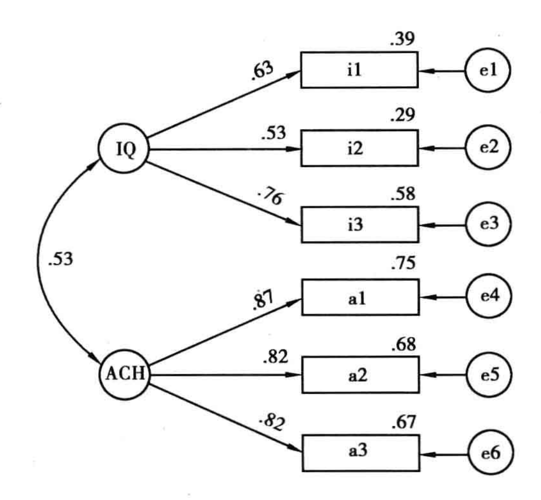


图 10-32　未受限模型的结果:标准化结果

表 10-10　IQ 与 ACH 相关系数估计值

			Estimate
IQ	<-->	ACH	.5263

由表 10-11 中"Par Change"(即 Expected parameter change)知:IO 与 ACH 间的参数释放估计时,其实际的参数值预估为:1 + (-0.3951) =0.6049,与实际的参数值(0.5263)略有差异。实际的新参数估计值约等于旧有的路径系数 + Par Change。通常旧有的路径系数设定为 0 时,实际的参数值略等于 Par Change;本例旧有的路径系数设定为 1,所以实际的相关系数参数值预估约等于:1 + (-0.3951) = 0.6049。前述非受限与受限模型的 χ^2 差异检验结果摘要如表 10-12。

表 10-11　IQ 与 ACH 间相关系数预估的改变值

			M. I.	Par Change
IQ	<-->	ACH	34.6067	-.3951

根据表 10-12 嵌套模型间的比较知,这两个受限与非受限模型间的卡方差异值(57.8768 -3.9181 = 53.9587)达显著差异(df = 1, P<0.001, TLI = 0.2621),表示这两个构念间具区辨性。

※实例应用

表 10-13 是另一区辨效度的研究实例,取自 Mesquita,Anand 与 Brush(2005)的论文《Knowledge Transfer in Vertical Alliances: Advantages for Whom?》。因为区辨效度的 SEM 分析,研究者一次仅能进行一对构念的检验,故表 10-13 中共进行了五次 SEM 区辨效度分析。表 10-13 中最后一栏的卡方差异值均大于 α =0.05 的显著水平 3.85(df = 1),故"Supplier knowledge acquisition effect"与其他潜在变量间均具有区辨效度。

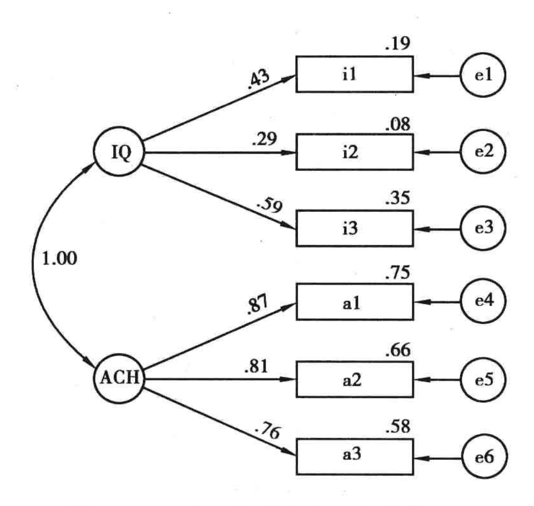


图 10-33 受限模型的结果:标准化结果

表 10-12 Amos 嵌套模型间的比较分析报表

Model	NPAR	CMIN	DF	P	CMIN/DF
The Unconstrained Model	12	3.9181	9	.9167	.4353
The Constrained Model	11	57.8768	10	.0000	5.7877

Assuming model The Unconstrained Model to be correct:

Model	DF	CMIN	P	NFI Delta-1	IFI Delta-2	RFI rho-1	TLI rho-2
The Constrained Model	1	53.9587	.0000	.1679	.1728	.2499	.2621

(五)多项特质—多项方法

多项特质—多项方法(multitrait-multimethod approach,简称 MTMM 方法)。此种方法由 Campbell 与 Fiske(1959)所创用,其检验较繁琐,研究者必须使用多重构念与多种方法去搜集数据,但它可同时取得区辨效度(discriminant validity 与聚敛效度(convergent validity)的证据,并可检验"方法因素"的干扰程度。以表 10-14 中"知识经济指标的建构效度研究"为例(陈智凯,2005),就研究设计来看,本研究同时涉及特质因素及方法因素。其中,特质因素包含知识经济的五大指标:企业经营环境、创新系统、人力资源发展、信息通讯科技、绩效指标;方法因素包含三个经济发展的研究机构:经济合作暨发展组织(OECD)、亚太经济合作组织(APEC)、世界银行(WB)。在表 10-14 MTMM 矩阵的主斜对角线中,相同方法—相同特质的相关系数均为 1,研究者可以信度系数替代之,理论上此对角线上的系数在 MTMM 矩阵中应为最高。其次,在不同方法—相同特质的对角线上的效度系数(0.488 ~0.967),为聚敛效度的指标,因此对角线的效度系数均应显著大于0,故知识经济衡量指标在不同方法上具有良好的聚敛效度。至于不同方法—相同特质的对角线效度系数(0.488 ~0.967),应显著高于在不同方法—不同特质的相关系数,在本例中除了"创新系统"测量指标的系数:A2W2(0.496)、O2W2(0.488)之外,其余测量

指标均具有良好的区辨效度。最后值得一提的是,在不同方法—不同特质对角线上的效度系数,理论上来说,其相关系数应最低,在本例中有些异常案例。至于相同方法—不同特质的区块中的相关系数,假如偏高极可能反映出测量结果受到"方法因素"的干扰,或特质间的区辨效度不佳。例如:在相同方法—不同特质区块中的相关系数:A3 - A4 = 0.816,O3 - O4 = 0.619,W3 - W4 = 0.809,这些相关系数似乎偏高,可能是使用相同方法,或特质间的区辨效度不佳所致。

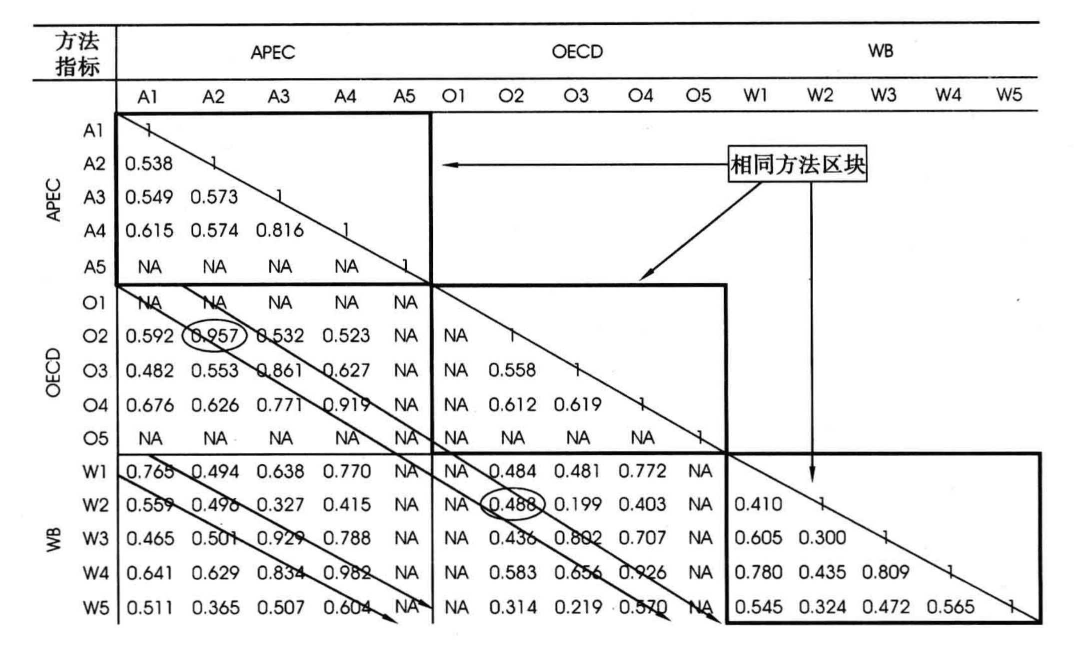
表 10-13　嵌套模型间的卡方差异检验

协方差	潜在变量		χ^2		差异值(df =1)大于 3.85 即表示达显著水平(α =0.05)
			限制模型(df =81)	非限制模型(df =80)	
CF1F2	Supplier Knowledge Acquisition Efforts	Joint Buyer Supplier Knowledge Acquisition Efforts	289.4	233.9	55.5
CF1F3	Supplier Knowledge Acquisition Efforts	Supplier Dyadic-Specific Assets & Capabilities	290.4	233.9	56.9
CF1F4	Supplier Knowledge Acquisition Efforts	Buyer-Supplier Relational Governance	297.5	233.9	63.6
CF1F5	Supplier Knowledge Acquisition Efforts	Re-deployable Performance	238.6	233.9	4.7
CF1F6	Supplier Knowledge Acquisition Efforts	Relational Performance	242.1	233.9	8.2

※实例应用

前述 MTMM 分析,因同时涉及特质因素及方法因素在实务上较难实现,其实研究者亦可将方法因素舍去,而简化如图 10-34 中之"问题解决构念"设计。理论上,我们假设"问题解决构念"与"事实知识构念"两者独立无关,每一构念下各拥有三个代表性题目反映该构念。经过数据分析之后,我们发现各个构念之间的关系不高,亦即这两个构念间具有区辨效度(参见图 10-34 下半部右上角与左下角相关矩阵),而且也发现每一构念下的题目间的相关矩阵,反映出这些同一建构的题目均聚敛到同一构念,亦即具有聚敛效度(参见图 10-34 下半部左上角与右下角相关矩阵)。研究者根据图 10-34 中的区辨力与聚敛效度的证据,即可认为"问题解决"具有建构效度。

表 10-14　知识经济衡量指标的 MTMM 矩阵(N = 100)



方法 指标		APEC					OECD					WB				
		A1	A2	A3	A4	A5	O1	O2	O3	O4	O5	W1	W2	W3	W4	W5
APEC	A1	1														
	A2	0.538	1													
	A3	0.549	0.573	1												
	A4	0.615	0.574	0.816	1											
	A5	NA	NA	NA	NA	1										
OECD	O1	NA	NA	NA	NA	NA										
	O2	0.592	0.957	0.532	0.523	NA	NA	1								
	O3	0.482	0.553	0.861	0.627	NA	NA	0.558	1							
	O4	0.676	0.626	0.771	0.919	NA	NA	0.612	0.619	1						
	O5	NA	NA	NA	NA	NA	NA	NA	NA	NA	1					
WB	W1	0.765	0.494	0.638	0.770	NA	NA	0.484	0.481	0.772	NA					
	W2	0.559	0.496	0.327	0.415	NA	NA	0.488	0.199	0.403	NA	0.410	1			
	W3	0.465	0.501	0.929	0.788	NA	NA	0.436	0.802	0.707	NA	0.605	0.300	1		
	W4	0.641	0.629	0.834	0.982	NA	NA	0.583	0.656	0.926	NA	0.780	0.435	0.809	1	
	W5	0.511	0.365	0.507	0.604	NA	NA	0.314	0.219	0.570	NA	0.545	0.324	0.472	0.565	1

注:1. A 代表 APEC 方法;O 代表 OECD 方法;W 代表 WB 方法;1～5 代表衡量指标的五大特质;NA 代表相关系数不可得。A1～A5 表示以 APEC 方法得到衡量指标(measures),然而 APEC 未纳入第 1 项衡量指标,故 A1 不可得,O1～O5 表示以 OECD 方法得到衡量指标,同样 OECD 未纳入第 1 及第 5 项衡量指标,故 O1 及 O5 不可得。

2. A1O1W1 表示企业经营环境衡量指标(Trait 1):A2O2W2 表示创新系统衡量指标(Trait 2);A3O3W3 表示人力资源发展衡量指标(Trait 3);A4O4W4 表示信息通信科技衡量指标(Traint 4);A5O5W5 表示绩效衡量指标(Trait 5),唯 A1,O1,O5 不可得。

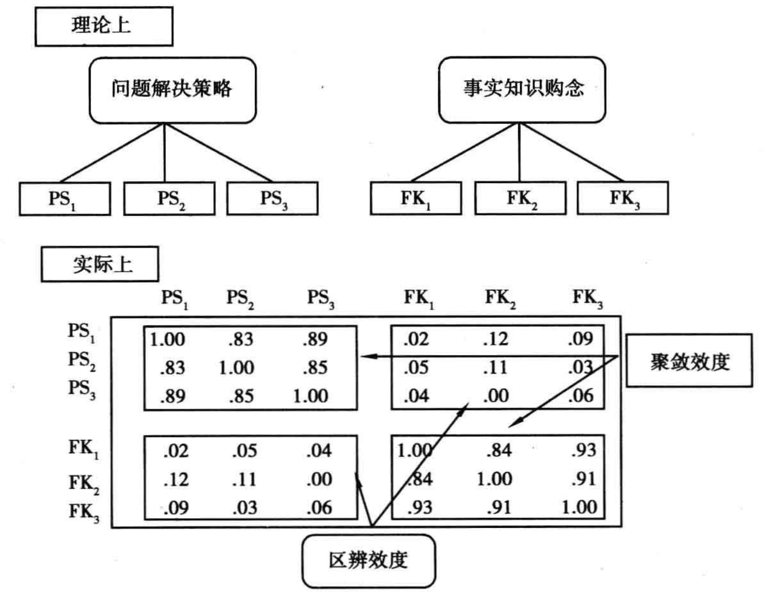


图 10-34　MTMM 的简化相关矩阵(无方法因素):区辨效度与聚敛效度

资料来源:修订自 http://www.msu.edu/course/psy/395/deshon/Lectures/Lectures%20Test%202/Lecture%2005%20-%20Validity.ppt。

四、聚敛效度

聚敛效度(convergent validity)的评估可以通过内部一致性(internal consistency)与效标关联效度(criterion validity)进行。内部一致性指各指标间应具有密切相关,或同因素内各题目间的相关要高,常用统计指标有建构信度(composite reliability)、Cronbach α(通常须大于 0.80)、单因子结构分析(各指标无跨因子的现象)、AVE(average variance extracted,通常须大于 0.50);效标关联效度,指相关构念间的密切关系,研究者尤其重视效标的表现,常用的统计指标有同时效度(效标可同时取得)与预测效度指标(效标须在未来才能取得),通常都以相关系数检验。例如,Goo,Kim 与 Cho(2006)的论文:《Structure of service level agreements (SLA) in IT outsourcing: The construct and its measurement》中,为了提供各量表的聚敛效度的证据,将各量表的构念信度、AVE 与各题目的因素负荷量(含 t 统计量,Amos 中称为 CR 值,用来检验因素负荷量是否为0)摘要如表 10-15 所示。因为各指标在各建构上的因素负荷量均较高,且能解释大部分相关潜在构念的变异量,各建构信度均大于0.87,各建构的 AVE 均大于0.69,显现出这些量表内部的题目均具有不错的聚敛效度。

表 10-15 聚敛效度检验:建构信度、AVE 与因素负荷量

Constructs	#of Items	Composite Reliability^	Average Variance Extracted	Loading (t-Statistice)*
Service Level Objectives	3	0.87	0.69	0.83(19.01), 0.84(17.33), 0.81(12.47)
Process Ownership Plan	3	0.92	0.80	0.94(62.93), 0.89(32.61), 0.85(22.68)
Service Level Contents	3	0.91	0.77	0.90(50.49), 0.89(35.26), 0.83(20.31)
Future Demand Mgmt Plan	3	0.93	0.81	0.89(35.81), 0.90(37.28), 0.91(39.86)
Anticipated Change Plan	3	0.92	0.80	0.90(42.75), 0.89(29.37), 0.90(35.02)
Innovation Plan	3	0.90	0.75	0.75(8.94), 0.90(37.87), 0.92(52.41)
Feedback Plan	3	0.94	0.84	0.90(43.50), 0.91(19.15), 0.93(52.56)
Communication Plan	3	0.93	0.81	0.88(29.53), 0.91(29.34), 0.91(43.10)
Measurement charter	3	0.93	0.82	0.90(37.49), 0.89(33.16), 0.92(35.83)
Conflict Arbitration charter	3	0.89	0.73	0.82(14.18), 0.84(28.00), 0.90(28.37)
Enforcement Plan	3	0.91	0.76	0.87(28.70), 0.91(36.34), 0.84(14.77)

^ The composite reliability scores were calculated with the formual prescribed by Fornell and Larcker(1981).

* $p < .001$.

同时效度分析的实例,请参看 Bagozzi(1981)的论文,该研究旨在利用捐血者的捐血态度,比较通过语意区分法(semantic differential measures)与预期价值法(expectancy-value measures)所编制出来的工具效度。文中检验相关工具的同时效度与区辨效度,其相关矩阵,如表 10-16 所示。由该表知,Aact(对于捐血的态度)与 EV1、EV2、EV3(预期价值态度)的相关系数达到 0.05 以上的显著水平,支持了同时效度;而 Aact(对于捐血的态度)与 EV1、EV2、EV3(预期价值态度)的相关系数显著小于 1,或 EV1、EV2、EV3 间的相关系数显著小于 1(可利用 2 × SE 检验之,$\alpha = 0.05$;例如:0.559 ± 2 × 0.102、0.787 ± 2 × 0.073,其置信区间并未包含 1),亦显现出这些量表建构内的题目均具有不错的区辨效度。

表 10-16 同时效度与区辨效度的相关矩阵

Intercorrelation Mattrices Among Attitudinal Constructs for Tests of Concurrent and Discriminant Validity

A. *past donorsa*	Aact	EV1	EV2	EV3
Aact	1.000			
EV1	.605(.088)[a]	1.000		
EV2	.280(.118)	.559(.102)	1.000	
EV3	.354(.114)	.429(.116)	.787(.078)	1.000
B. *Current donors*[b]	Aact	EV1	EV2	EV3
Aact	1.000			
EV1	.413(.085)	1.000		
EV2	.527(.074)	.670(.064)	1.000	
EV3	.497(.078)	.786(.052)	.785(.050)	1.000

注:The first column in each matrix represents the correlations between the samantic differential attitude toward the act construst and the respective dimensions of the expectancy-value model. a. n = 81 b. n = 127 c. standara errors in parentcses

研究者亦可利用图 10-35 的路径模型,检验相关的路径系数(φ_{21}、φ_{31}、φ_{41})是否显著大于 0(在 Amos 中可查看各参数估计值的 S. E.,C. R. 与 p 值)。

预测效度的分析,研究者可通过如表 10-17 的相关矩阵(取自 Bagozzi, 1981),进行回归系数的显著性检验(可利用 2 × SE 检验,$\alpha = 0.05$ 或利用 SPSS 的回归子程序进行显著性的检验),假如 0.95 的置信区间未包含 0,即显示该回归系数的估计值达 0.05 显著水平。表 10-17 的相关系数:0.301,0.447,0.368 的 0.95 置信区间分别为:0.067 ~ 0.535,0.219 ~ 0.675,0.136 ~ 0.60,其置信区间均未包含 0,反映出这些指标均具有良好的预测效度。当然,亦可通过 SEM 的途径(参见图 10-36),检验相关路径系数的显著性,Amos 会自动提供各参数估计值的 CR 值及 p 值。

表 10-17 预测效度的检验

Predkctive Validity of Expectancy-value Model: Attitude Predocting Behavioral Intentions

Regression Weight	Past Donors[a]	Current Donors[b]
γ_1	.301(.117)[c]	.374(.092)
γ_2	.447(.114)	.373(.092)
γ_3	.368(.116)	.536(.088)

a. n = 81 b. n = 127 c. standard errors in parentheses.

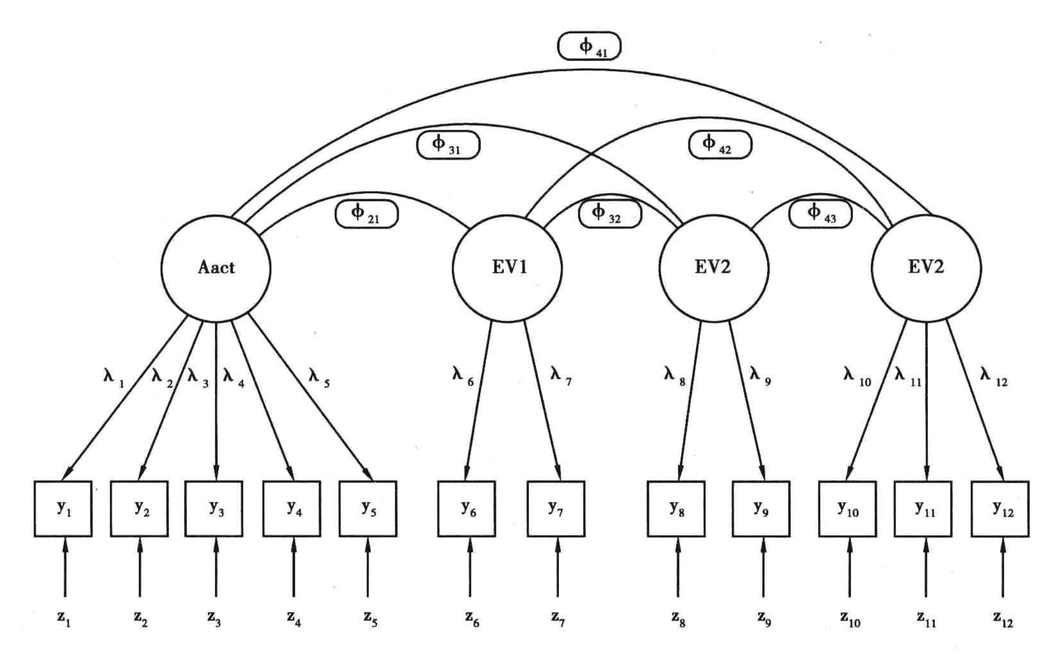


图 10-35　预测效度检验的路径设计

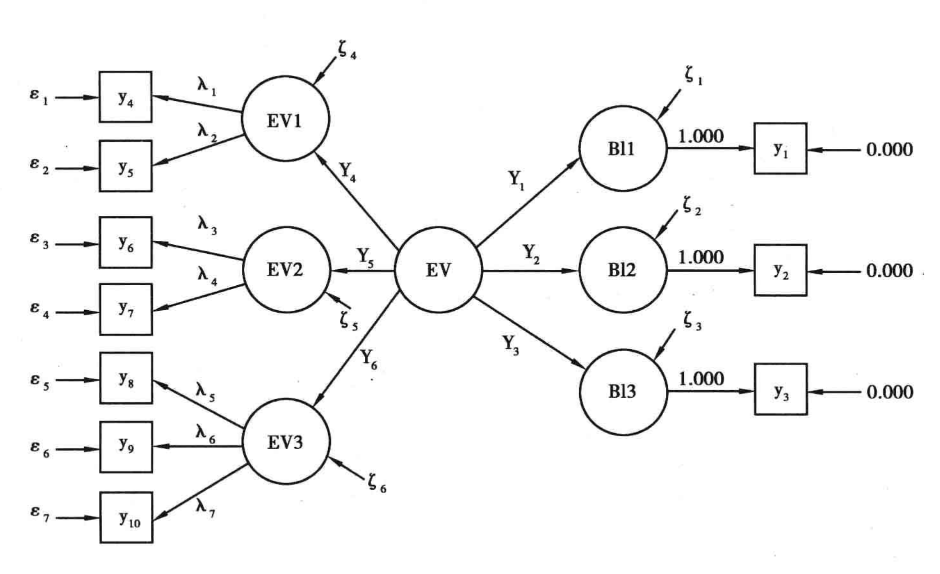


图 10-36　预测效度检验的路径设计

在此值得一提的是，内部一致性高或聚敛效度佳，有时可能是源于共同方法变异（common method variance，简称 CMV），亦即构念或指标间的相关是由于共同方法、相同受试者所导致，而非构念或指标特质所导致，这是测量编制者或研究者在解释时，应谨慎保守的地方。欲评估共同方法变异，研究者通常需要使用多重特质—多重方法（MTMM）的研究设计。图 10-37 中 SheMale 是 GSR 的聚敛效标；而 Neuroticism 是 GSR 的区辨效标。这三种特质分别用自陈法（Self-report）、同侪报告法（peer-report ）、日记（diary behavior report）测量。这些特值与方法的 MTMM 分析结果，摘要如图 10-37 所示（Aspelmeier,

2007)。请注意圈起来的相关系数,似乎因为使用相同方法而得到较高的相关系数。例如,GSRsr 与 Nsr(区辨效标)的相关理应偏低,但因均使用 Self-report 方法而出现偏高的构念相关(=0.55),当使用不同方法测量时,GSRsr 与 Npr 或 Ndr 间的相关就无此虚胖现象,这就致使研究者难以分辨到底构念间是真正相关还是共同方法所致。

	GSR sr	GSR pr	GSR dr	SHeMale sr	SHeMale pr	SHeMale dr	N sr	N pr	N dr
GSR sr	1.00								
GSR pr	0.80	1.00							
GSR dr	0.83	0.88	1.00						
SHeMale sr	0.45	0.06	0.16	1.00					
SHeMale pr	0.20	0.40	0.06	0.86	1.00				
SHeMale dr	0.15	0.15	0.48	0.85	0.84	1.00			
N sr	0.55	0.01	0.03	0.45	0.11	0.08	1.00		
Npr	0.10	0.50	-0.03	0.05	0.43	-0.03	0.85	1.00	
Ndr	-0.11	0.12	0.44	0.15	-0.09	0.55	0.86	0.90	1.00

图 10-37　共同方法变异的检验:MTMM 矩阵分析方法

注:取自 www. runet. edu/ ~ jaspelme/611/Spring-2006/Measurment%20Validity. pdf.

此外,研究者亦可通过验证性因素分析,针对 MTMM 矩阵加以检验,过去的研究者常利用以下四个模型来决定特质或方法变异量的存在程度(Cote & Buckley, 1987)。

1. 虚无模型:模型中仅有误差项,无特质或方法因素;
2. 特质模型:模型中仅有特质因素与误差项,无方法因素;
3. 方法模型:模型中仅有方法因素与误差项,无特质因素;
4. 特质与方法模型:模型中含有特质、方法因素与误差项,无特质与方法的交互相关。

假如模型 2 的适配度显著优于模型 1 的,且模型 4 的适配度显著优于模型 3 的适配度,研究者即能确定"特质因素"的存在;假如模型 3 的适配度显著优于模型 1 的,且模型 4 的适配度显著优于模型 2 的,研究者即能确定"方法因素"的存在。

目前最常用且较不会产生模型辨识问题的"方法因素"的检验方法有二(Podsakoff et al. , 2003; Meade, Watson & Kroustalis, 2007):一是 CTCM 模型;二是 CTCU 模型。

(一)CTCM 模型

特质相关—方法相关模型(the correlated-trait correlated-method model)至少需要三个特质与方法模型才能辨识。它除了充分利用 MTMM 的所有信息之外,还可以直接估计特质与方法的相对大小。通常假定特质与方法间独立无关。基本上,共同方法变异的评估是通过计算特质相关—方法相关模型与仅包含特质相关模型的适配度的差距。假如两个模型间的适配度差异未达统计显著水平,研究者即可宣称无共同方法变异存在,而偏爱较精简的特质相关模型。在标准化的情境下,各指标在相关方法因素上的因素负荷量的平方值,可以解释为方法变异的比率;而各指标在特质因素上的因素负荷量的平方值,可以解释为实际特质变异的比率。本模型在参数的估计过程中,如遇因素间缺乏区辨效度或各因素下的因素负荷量相等,常出现模型不可辨识或异常估计值的问题(如相关系数大于 1 或负的方差)。遇到此现象,请改用 CTCU 模型试试看。

(二)CTCU 模型

特质相关—误差相关模型(the correlated-trait correlated-uniqueness model)检验法在同一方法因素内的指标误差项允许具有相关,此模型比 CTCM 模型更容易收敛,而得到

正常解,但至少需要两个特质与方法模型才能辨识。基本上,共同方法变异的评估是通过计算特质相关—误差相关模型与仅包含特质相关模型的适配度的差距。假如两个模型间的适配度差异达统计显著水平,研究者即可宣称共同方法变异的存在。由于 CTCU 模型未包含方法因素,指标间的相关会反映在特质间相关系数及因素负荷量上,因而相比 CTCM 模型特质间相关系数及因素负荷量较高。但一般来说,CTCU 模型会比仅包含特质相关模型的估计值来得低(Meade, Watson & Kroustalis, 2007)。由此观之,"共同方法变异"的存在通常会导致特质间相关系数及因素负荷量的虚胖。此种虚胖现象将在以下实例中加以说明。

※实例解说

Manson, Levine & Brannic(2000)为探究工作评定量表的建构效度,乃要求现职工作人员与管理人员,利用相对性评估与绝对性评估方法,评定工作的四个重要面向:重要性、错误的严重性、学习难度、所需时间。因此,本研究涉及以下三种评估方法:

M1:现职人员相对性评估;

M2:现职人员绝对性评估;

M3:管理人员相对性评估;

及涉及以下四个工作的特质:

IM:重要性;

CE:错误的严重性;

DL:学习难度;

DS:所需时间。

他们的研究所得的数据,摘述如图 10-38 所示,这是工作分析量表分数的 MTMM 数据相关矩阵,建置于 SPSS 数据编辑器中。

MANSON 2000.sav - SPSS Data Editor

25 : M2DL

	rowtype_	varname_	M1IM	M1CE	M1DL	M1DS	M2IM	M2CE	M2DL	M2DS	M3IM	M3CE	M3DL	M3DS
1	N		80.00	80.00	80.00	80.00	80.00	80.00	80.00	80.00	80.00	80.00	80.00	80.00
2	CORR	M1IM	1.00	.	.	.	.	.	.	.	.	.	.	.
3	CORR	M1CE	.90	1.00	.	.	.	.	.	.	.	.	.	.
4	CORR	M1DL	.54	.43	1.00	.	.	.	.	.	.	.	.	.
5	CORR	M1DS	.18	.18	.13	1.00	.	.	.	.	.	.	.	.
6	CORR	M2IM	.81	.77	.57	.29	1.00	.	.	.	.	.	.	.
7	CORR	M2CE	.74	.75	.27	.08	.75	1.00	.	.	.	.	.	.
8	CORR	M2DL	.57	.47	.85	.11	.68	.44	1.00	.	.	.	.	.
9	CORR	M2DS	.17	.17	.06	.66	.36	.36	.16	1.00	.	.	.	.
10	CORR	M3IM	.84	.83	.48	.12	.85	.71	.54	.16	1.00	.	.	.
11	CORR	M3CE	.86	.84	.48	.16	.75	.73	.50	.20	.88	1.00	.	.
12	CORR	M3DL	.55	.44	.90	.01	.56	.33	.85	.05	.54	.52	1.00	.
13	CORR	M3DS	.12	.15	.19	.78	.32	.08	.16	.60	.24	.17	.11	1.00
14	STDDEV		.88	.89	.99	.73	.98	1.20	1.15	.78	1.08	1.12	1.20	.88
15	MEAN		4.31	4.04	3.36	2.87	3.82	3.72	4.15	1.93	4.66	4.07	3.34	3.06

Data View / Variable View

SPSS Processor is ready

图 10-38　工作分析量表分数的 MTMM 数据矩阵

资料来源:Manson, Levine, Brannick(2000)

以下为了检验"方法因素"的存在,将通过 CFA 的 CTCM 和 CTCU 模型进行检验。

※ CTCM 模型

图 10-39 是传统 SEM 分析模型,假定误差项间独立无关,所得的参数估计值如图 10-39 中的路径系数。

由图 10-39 中的路径设计知,此模型仅包含特质且误差项间独立无关,本 CT 模型的适配度并不佳(162.928, df = 48, P = 0.0000)。模型中,T1 表示所需时间,T2 表示学习难度,T3 表示错误的严重性,T4 表示重要性。在各个工作特质下的指标间,其聚敛效度

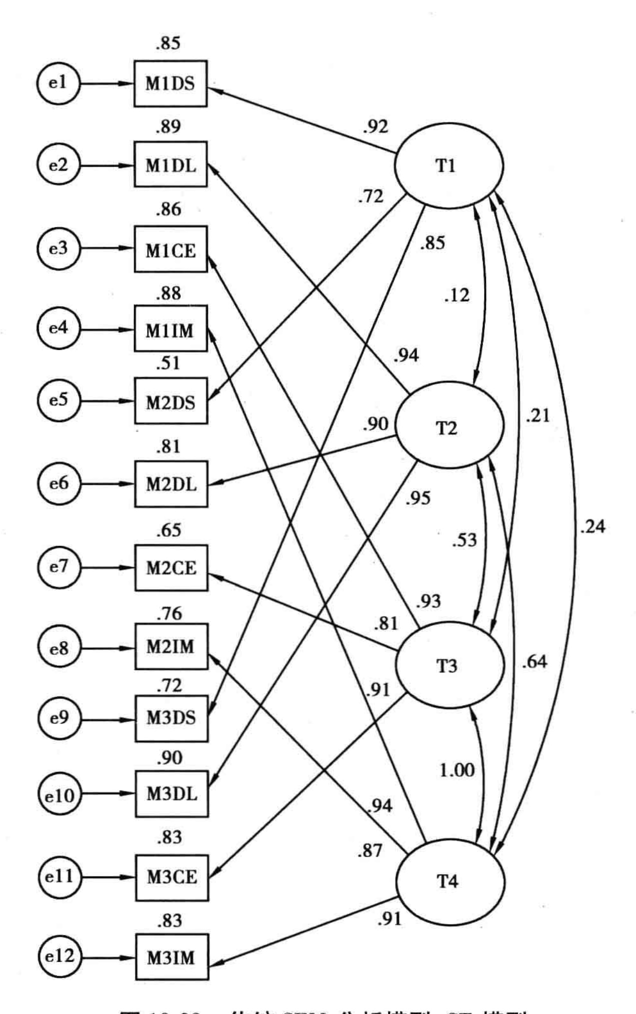


图 10-39 传统 SEM 分析模型:CT 模型

均甚佳(解释量均超过 50% 以上)。例如,T1 的 AVE = 0.66,T2 的 AVE = 0.87,T3 的 AVE = 0.78,T4 的 AVE = 0.82。但此高聚敛效度是否因“方法因素”所致,尚待以下 CTCM 模型的检验,参见图 10-40。

比较图 10-39 与图 10-40 中相对应的参数可知,CTCM 中参数估计值要比 CT 中来得小。请注意图 10-40 中 T3 与 T4 间的相关系数超过 1,这是不合理的参数估计值,似乎反映了这两个特质间缺乏区辨效度。CTCM 模型的分析结果如表 10-18 所示。

本模型虽然可以辨识而计算出 χ^2 值,但却产生非正定矩阵的警讯:This solution is not admissible,反映出有些估计参数超出合理范围,请参见图 10-41 内涵,该模型可能界定错误或有多重共线性。原本不应继续分析下去,但为进行实例解说 CMV 的影响情形,笔者仍继续分析下去。

表 10-18 CTCM 模型的分析结果

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	45	62.5730	33	.0014	1.8962

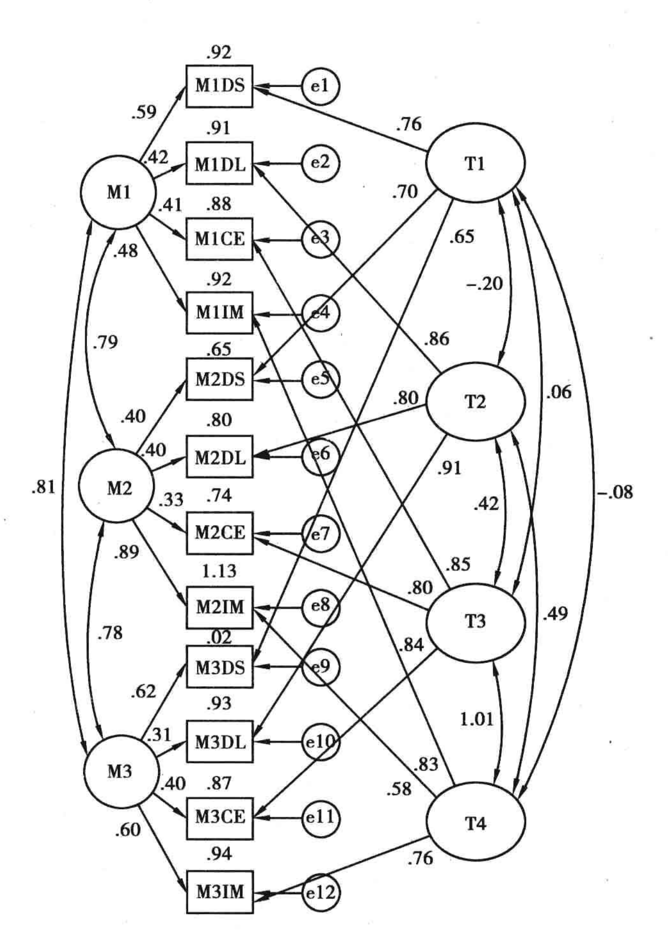
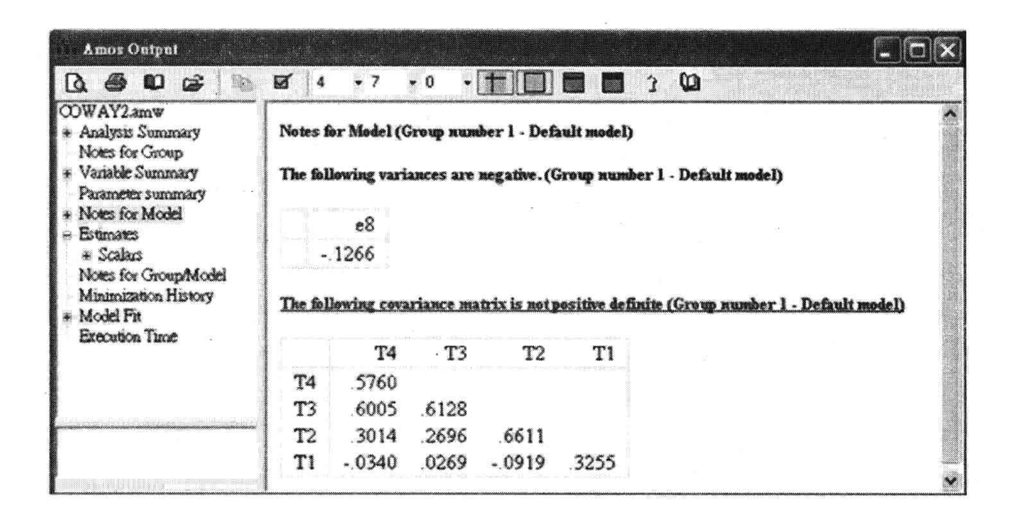


图 10-40　CTCM 模型：标准化参数值



	T4	T3	T2	T1
T4	.5760			
T3	.6005	.6128		
T2	.3014	.2696	.6611	
T1	-.0340	.0269	-.0919	.3255

图 10-41

表 10-19　CTCM 模型中各种变异源的分析

测量方法	方法变异源	特质变异源	测量误差变异源
M1	.35	.57	.08
	.17	.74	.09
	.16	.72	.12
	.23	.69	.08
M1 平均变异源	.23	.68	.09
M2	.16	.49	.35
	.16	.64	.20
	.10	.64	.26
	.79	.34	-.13 *
M2 平均变异源	.30	.53	.17
M3	.38	.42	.20
	.10	.83	.07
	.16	.71	.13
	.36	.58	.06
M3 平均变异源	.25	.64	.11
总平均变异源	.26	.62	.12

* 因非正定矩阵产生异常值。

表 10-19 中数据根据标准化的路径参数平方值计算而得,就整体来看,真正特质的总平均变异源比率约在 0.62,方法的总平均变异源比率约在 0.26,而测量误差的总平均变异源比率约在 0.12。CTCM 模型的适配度比 CT 模型更佳,χ^2 差异(162.93 - 62.57, df = 48 - 33) 检验结果达显著水平,反映出"共同方法变异"的存在。以下将以另一 CTCU 模型,进行比较分析,模型设计如图 10-42 所示。为了进行误差项参数的限制,将误差项命名为:c1 ~ c18。

※CTCU 模型

笔者为了比较 CTCM 与 CT 模型的适配性,利用模型管理窗口建立了两个分析模型:Correlated model 与 Uncorrelated model。Uncorrelated 模型的参数限制如图 10-43 所示。

由于 CTCU 模型未包含方法因素,指标间的相关会反映在特质间相关系数及因素负荷量上,因而导致比 CTCM 模型中的特质间相关系数及因素负荷量来得高。通常,误差项间的高相关亦反映出"共同方法变异的存在"。请检查图 10-44 的 CTCU 模型中误差项的标准化参数值的大小。注意图 10-44 中 T3 与 T4 间的相关系数达 0.98,似乎亦反映出这两个特质间缺乏区辨效度。

由表 10-20 知,CTCU 模型的适配度比误差项独立模型更佳,其 χ^2 差异(162.93 - 71.11)检验结果如表 10-21 所示:

表 10-21 的统计检验结果,亦反映出 CTCU 模型的适配度比误差项独立模型更佳,达 $\alpha = 0.05$ 的显著水平($p = 0.0000$),"方法效应"的确存在。换言之,该研究测量工具的效度已受到了威胁。

表 10-20　CT 模型与 CTCU 模型的适配度检验

Model	NPAR	CMIN	DF	P	CMIN/DF
Uncorrelated model	30	162.9282	48	.0000	3.3943
Correlated model	48	71.1144	30	.0000	2.3705

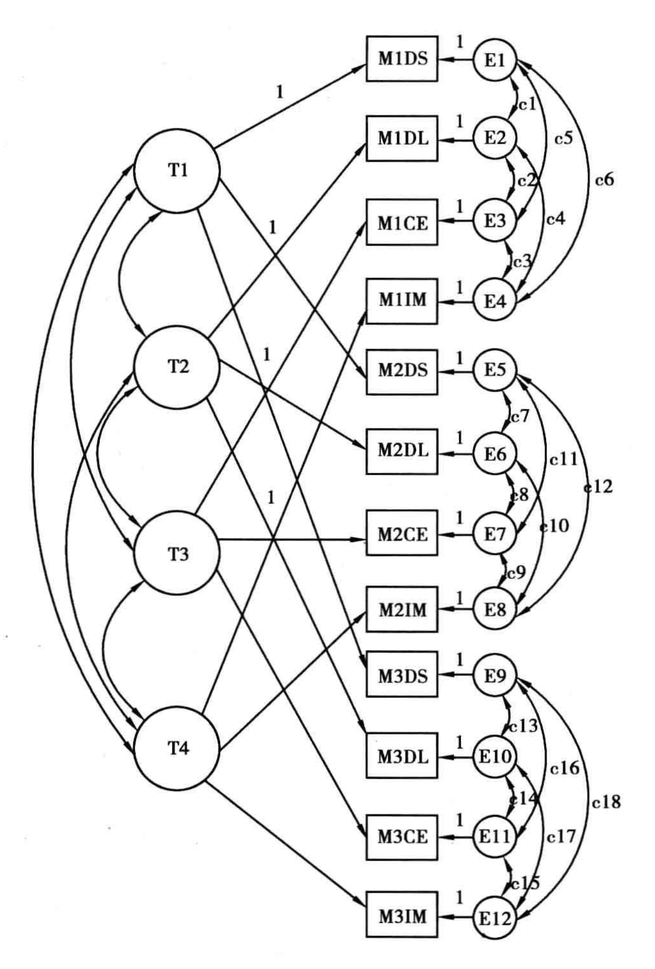


图 10-42　CTCU 模型

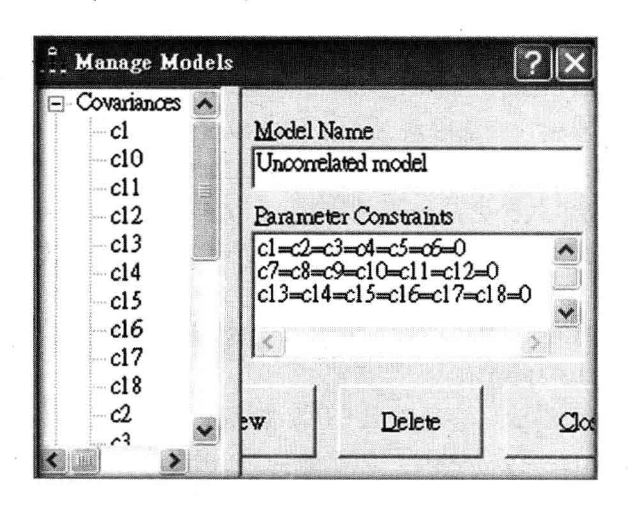


图 10-43　CTCU 模型的参数限制

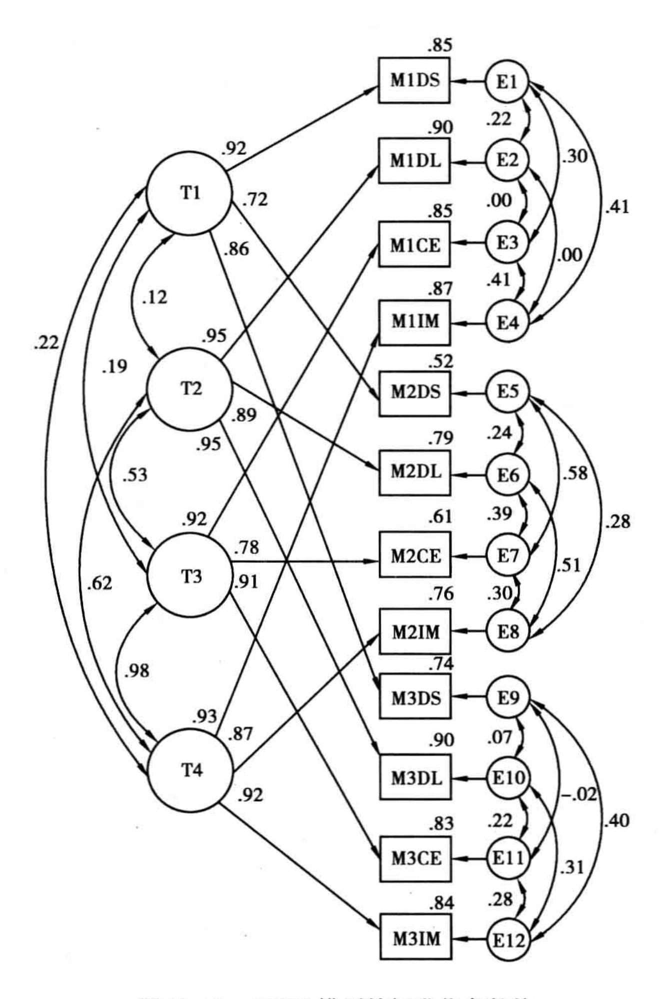


图 10-44 CTCU 模型的标准化参数值

表 10-21 CT 模型与 CTCU 模型的比较(Assuming model Correlated model to be correct)

Model	DF	CMIN	P
Uncorrelated model	18	91.8138	.0000

近年来,Lindell 与 Whitney(2001)在问卷中设置一个与实质变量无关的标竿变量(marker variable),再利用净相关技术进行共同方法变异量的估计。Hartman,Williams 与 Cavazotte(2002)列举了 Lindell 与 Whitney 氏方法的八大应用步骤,并举实例加以说明,有兴趣的读者可以参考。运用此法简单易行,但其最大的难题在于找出与理论建构无关的标竿变量。

五、构念网络效度

构念网络效度(Nomological Validity),亦称为"建构网络效度",是建构效度中的一种,指待研究的建构在相关理论网络系统内,待研究的构念与其他同时存在的外在效标,及与多个前因或后果变量间的关系程度(Cronbach & Meehl, 1955; Bagozzi, 1981; Churchill, 1995)。因此,构念网络效度的建立至少要有以下几个要素:两个以上的建构、

建构间的理论假说、能够加以测量的实征指标、指标变量与建构间的连接关系。研究者可根据理论或测量研究假设，建立出构念与构念之间，及构念与外显行为间的预期的系统网络关系，并绘出相关的路径图，并通过 SEM 进行构念网络效度的分析，以检验各潜在变量的 R^2（或可解释的百分比，即内在效度）与预期路径关系的关系方向与强度（即外在效度）。图 10-45 是美国顾客满意度指标的建构网络效度的检验模型，其中顾客满意度指标由前导变量“期待质量”、“感受价值”与“感受质量”等三个潜在构念组成，而“顾客抱怨”与“顾客忠诚”等两个潜在构念是“顾客满意”的结果变量。各潜在变量间的网络关系，如此模型中的因果关系（单箭头）或共变关系（双箭头）所示。由此预期的网络关系可知，顾客期待质量与感受质量互为共变，顾客期待质量会影响感受价值与顾客满意度，感受质量会影响感受价值与顾客满意度，而顾客满意度则会影响顾客对管理与人事上的抱怨、顾客的再购意愿、对价格的容忍，顾客的抱怨亦会直接影响顾客的忠诚度。

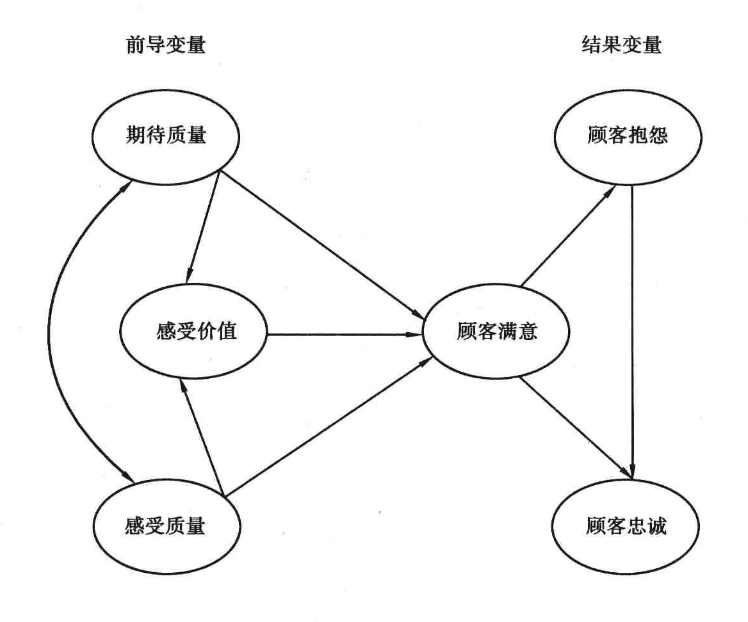


图 10-45　顾客满意度的构念网路效度检验模型

资料来源：Anderson 与 Fornell (2000)

※实例解说

以 Bagozzi(1981) 的捐血态度预期价值量表的效度研究为例，捐血的态度（EV）与个人的捐血动机（BI）、过去的捐血次数（PB）、社会看法（SNB）、个人看法（PNB）间具有密切关系，其检验的 Amos 构念网络如图 10-46 所示：

图 10-46 的模型经 SEM 统计分析后，就过去或目前的捐血者（Past donors or Current donors）而言，该模型的适配度均甚佳（$\chi^2 = 510.05$, p =0.51, df = 52; $\chi^2 = 57.79$, p = 0.30, df = 53）。表 10-22 摘要该构念网络效度模型中的相关参数估计值。

由表 10-22 的参数估计值及标准误可知，就过去的 81 位捐血者（past donors）而言，预期价值变量（EV）与个人的捐血动机（BI1 ~ BI3）具有正相关，但未达 0.05 的显著水平；就目前的 127 位捐血者（current donors）而言，预期价值变量（EV）与个人的捐血动机

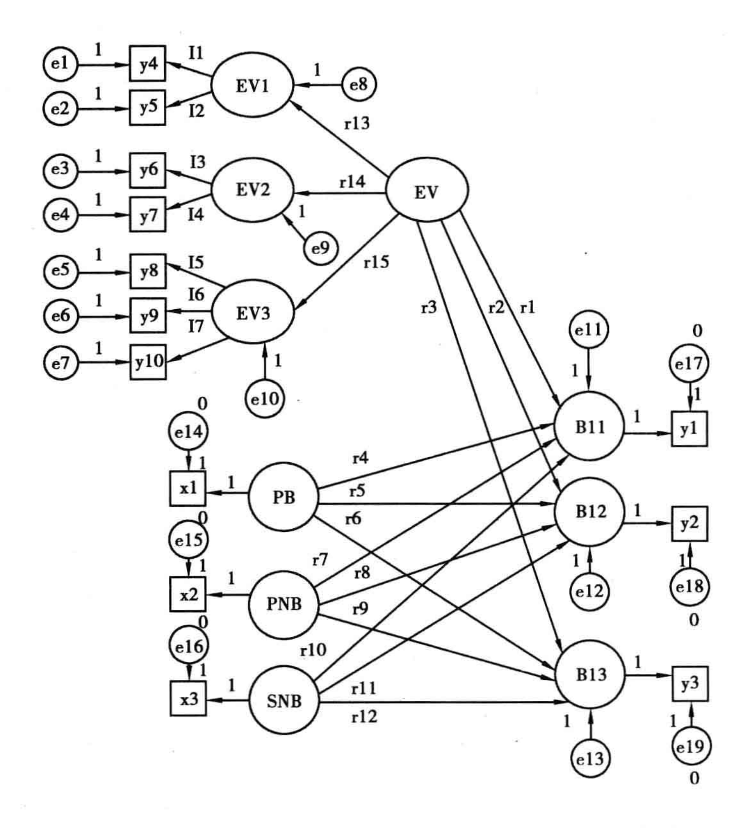


图 10-46　态度预期价值模型的建构网络效度检验模型

(BI1 ~ BI3)具有显著正相关(达 0.05 的显著水平)，因为相关系数均落在两个标准误的置信区间之外。另外，过去的捐血次数(PB)对于过去或目前个人的捐血动机(BI1 ~ BI3)具有显著的预测力，而社会看法(SNB)与个人看法(PNB)则无显著的预测力。因此，Bagozzi(1981)的构念网络效度并未获得完全支持。

表 10-22　建构网络效度的相关参数的估计值摘要表

predictor	Past Donors[a]			Current Donors[b]		
	BI1	BI2	BI3	BI1	BI2	BI3
	Parameter Value	Parameter Value	Parameter Value	Parameter Value	Parameter Value	Parameter Value
EV	.102(.110)[c]	.180(.098)	.163(.108)	.341(.099)	.222(.090)	.424(.089)
PB	.409(.111)	.542(.098)	.492(.108)	.108(.094)	.393(.086)	.289(.088)
PNB	.162(.104)	.102(.091)	−.061(.101)	.018(.089)	.074(.081)	.050(.078)
SNB	−.052(.107)	.084(.094)	.098(.104)	−.090(.087)	−.027(.080)	.018(.077)

资料来源：Symbols are defined in text.　a. n = 81　b. n = 127　c. standard errors in parentheses.

六、建构效度

建构效度(construct validity)，亦称为构念效度，用以检验所搜集到的数据是否能支持所宣称的理论假设。建构效度又称为因素效度(factorial validity)，由 Cronbach 与 Meehl

(1955)所创用,依照 Cronbach 与 Meehl 的定义,建构效度旨在探究下列问题:

有哪些构念可以说明测量分数中的变异量?

因此,最常用的探究途径是进行维度分析或因素结构分析,以探索或验证量表内容的因素结构,常用的统计方法为探索性(用在项目分析阶段)与验证性(用在效度检验阶段)的因素分析。近年来,一般研究者都认为效度乃是单一的概念(a unitary concept),因此内容效度、效标关联效度与构念网络效度均视为建构效度的一员(Shepard, 1993; Straub, Boudreau & Gefen, 2004)。因而学界乃出现效度只有一种,但求证方法是多元的呼声。依此,建构效度的检验途径当然很多元,例如:内部结构分析、群体平均数差异法、相关分析法、因素分析法、MTMM 与构念网络效度。此外,Western ,Rosenthal(2003)与 Taras , Kline(2007)又分别提出了建构效度系数的简单统计方法:Western 与 Rosenthal 的 $R_{alerting\text{-}CV}$ 可以进行不同研究间或不同概念间的比较;Taras 的 D_h 系数,可以评估题目的建构效度(item construct validity),以进行项目分析及建构效度分析。兹依因素结构分析与建构效度系数的研究实例,分别简介如下:

※因素结构分析实例

Kane 与 Krenzer(2006)根据过去的成人 IQ 理论与文献,提出两个 IQ 竞争模型,以验证测量分数变异量的 IQ 构念因素。第一个提议模型为理论模型,是二阶的因素模型,第一阶因素为通用因素,第二阶因素则包含有语文理解、工作记忆、知觉组织与处理速度。第二个提议模型为临床应用模型,是三阶的因素模型,第一阶因素为通用因素,第二阶因素则包含语文智商与操作智商,第三阶因素则包含有语文理解、工作记忆、知觉组织与处理速度。研究者通过 CFA 验证与比较这两个竞争模型,分析结果如图 10-47 与 10-48 所示。

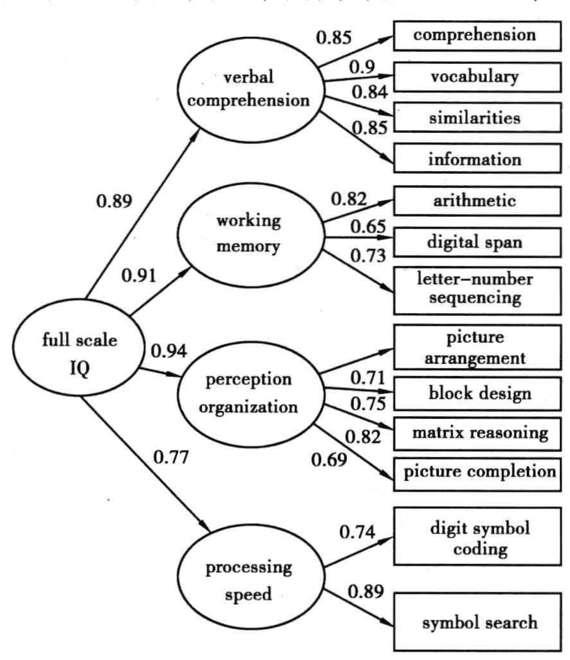


图 10-47 魏氏成人智力量表的理论模型与分析结果

资料来源:Kane 与 Krenze(2006)

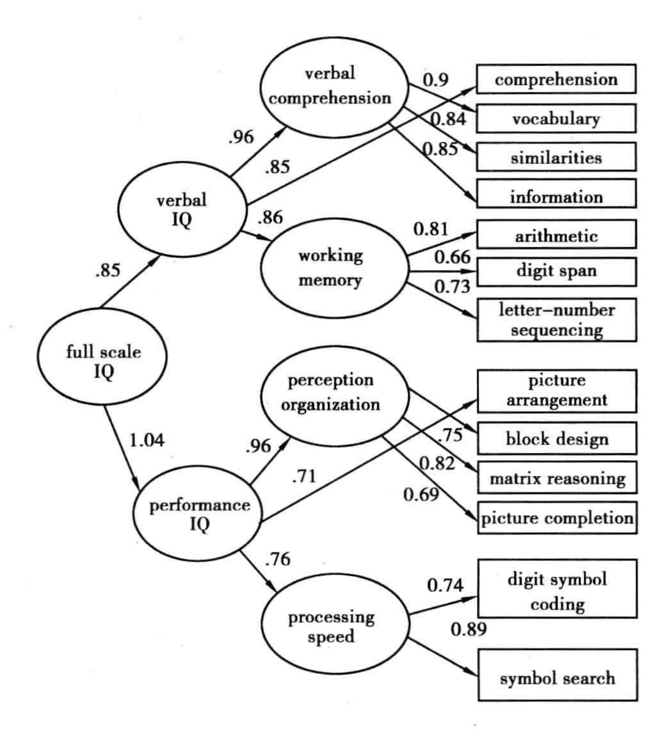


图 10-48　魏氏成人智力量表的临床模型与分析结果

资料来源:Kane 与 Krenzer(2006)

根据图 10-47 与 10-48 中魏氏成人智力量表的理论与临床模型的比较分析,相关参数的估计值,摘要如表 10-23 所示:

表 10-23　魏氏成人智力量表的理论与临床模型的比较

模型	χ^2	df	p	AGFI	TLI	RMSEA
理论	65.01	61	.34	.93	1.0	.02
临床	75.21	60	.09	.92	.99	.04
$\triangle\chi^2$ =10.11*, $\triangle$df = 1						

*p<.05.

表 10-23 中的适配度指标显示,两个提议模型的适配度均佳,但理论模型的适配度更好。即魏氏成人智力量表的理论模型的因素结构与实际数据较适配,也就是二阶的因素结构较能说明测量分数中的变异量。另外,研究者亦可计算理论模型中四个一阶因素的建构信度:0.9193,0.7791,0.8317,0.8009,反映出指标间的聚敛效度甚佳。

※建构效度系数实例

A. 建构效度系数 $R_{alerting\text{-}CV}$

为了使用单一的指标来量化建构效度,Western 与 Rosenthal(2003)提出了两个效果值指标:$R_{alerting\text{-}CV}$ 与 $R_{contrast}$。前者计算简单且易于解释,可作为建构效度的参考指标,笔者特引用文中实例介绍如下。有研究者开发了一个诊断青少年人格异常的新工具,其与 DIS-IV(Diagnostic and Statistical Manual of Mental Disorder, 4th ed.)中 10 个人格异常诊断的理论及实征相关,摘要如表 10-24。

表 10-24 内第二栏,是理论上的预测值,亦即研究者预期新量表分数与这 10 个人格

异常诊断分析的相关系数；第三栏，是新量表分数与十个人格异常评定结果间的实际相关，亦即研究者针对新量表分数与这十个人格异常评定结果的实征相关系数。为了符合正态分布的基本假设，研究者亦可将表 10-24 内第三栏的相关系数，进行自然对数的 Fisher 的 z 转换（第四栏）。接着，研究者即可利用第二、三或四栏的数据计算积差相关，亦即建构效度系数 $R_{alerting\text{-}CV}$（=0.9036），反映出该新量表具有不错的建构效度。本指标较适合于效标变量大于三个以上时。当只有两个效标变量时，$R_{alerting\text{-}CV}$ 将只会出现 ±1，指标过于粗略用途不大。研究者如需计算该指标的 Excel 程序，可至右列网址索取：http://www.psychsystems.net/lab/。图 10-49 是建构效度系数 $R_{alerting\text{-}CV}$ 的 Excel 程序的界面及相关的统计量。

表 10-24 新量表与 DIS-IV 中 10 个人格异常诊断的理论相关及实征关系

1	2	3	4
Personality Disorder	Predicted correlations	Ratings（r）	Ratings（Z_r）
Histrionic	.60	.55	.62
Borderline	.30	.51	.56
Dependent	.10	.20	.20
Antisocial	.00	-.06	-.06
Narcissistic	.00	.10	.10
Paranoid	-.10	-.04	-.04
Obsessive-compulsive	-.40	-.23	-.23
Avoidant	-.50	-.20	-.20
Schizoid	-.50	-.15	-.15
Schizotvpal	-.50	-.02	-.02

资料来源：修订自 Western 与 Rosenthal（2003）。

B. 项目分析 D_h 系数

Taras 与 Kline（2007）的 D_h 系数，最适合于小样本时进行建构效度分析，并可作为项目分析用。计算此系数，研究者亦须事先界定待研究构念与其他相关变量间的预期关系，以建立理论上的效标关联系数。接着，研究者计算待研究构念与其他相关变量间的相关系数，以建立实际的效标关联系数。最后利用以下公式计算 Taras 的 D_h 系数。

$$D_h = 1 - \sqrt{\frac{\sum_{1\sim k}(r_h - r_o)^2}{k}}$$

式中 r_h 代表某一题目与效标的理论相关系数，r_o 代表某一题目与效标的实得相关系数，k 代表效标变量数目。D_h 的最大限制乃是易受主观性的影响而偏高，例如，当实得相关系数偏低时，研究者故意选用偏低的理论相关效标，D_h 值则会偏高，可改以 D_m 来计算：

$$D_m = 1 - \sqrt{\frac{\sum_{1\sim k}(0.5_{hd} - r_o)^2}{k}},\quad D_m = 1 - \sqrt{\frac{\sum_{i=1}^{j}\sum_{j=1}^{k}(0.5 - r_o)^2}{j * k}}$$

式中 0.5_{hd} 常数代表正向理论关系，如与效标的关系为负向则改为 -0.5_{hd}。以 0.5 作为检验的参照点，主要是植基于：(1)此值过高表示区辨效度不佳，而此值过低表示效标不良；(2)可以进行不同研究间的比较。表 10-25 中的数据，是研究者为了建立建构效

QUANTIFYING CONSTRUCT VALIDITY

Enter data in blue cells

sample size	266	
median intercorr	0.113	

criteria scales	PREDICTED CORRS	Actual CORRS
1 Achievement striving	0.60	.55
2 Flexibility	0.30	.51
3 Friendliness	0.10	.20
4 Anger	0.00	-.06
5 Anxiety	0.00	.10
6 Assertiveness	-0.10	-.04

Results are in red cells

	r alerting - CV	.983
	r contrast - CV	1.000
95% CI around r contrast	lower	1.000
	higher	1.000
	Z contrast	13.757
	p	[illegible]
	tcontrast	10000000.000
95% CI around contrast	SE contrast	.067
	lower	.791
	higher	1.054

INTEGERS OF RAW LAMBDAS	Zr OBS CORRS
.70	.62
.40	.56
.20	.20
.10	-.06
.10	.10
.00	-.04
-.30	-.23
-.40	-.20
-.40	-.15
-.40	-.02

图 10-49　建构效度系数 $R_{alerting\text{-}CV}$ 的 Excel 程序界面

度,邀请六位大学教授对组织承诺量表中四个题目作答的数据。

由于样本数据过少,不适于进行因素分析,研究者于是进一步取得表 10-26 中六位教授在八个效标变量上的反应数据。

表 10-25　六位教授在组织承诺量表中的反应数据

Respondents	Item1	Item2	Item3	Item4
Professor 1	5	6	6	5
Professor 2	5	4	6	4
Professor 3	4	5	7	5
Professor 4	5	5	4	3
Professor 5	5	5	6	4
Professor 6	3	6	5	5

资料来源:取自 Taras 与 Kline(2007)。

表 10-26　六位教授在八个效标变量上的反应数据

Respondents	Exogenous Theoretically Relevant Variables							
	UR	JS	PS	Coll	Age	Rel	SJ	SJS
Professor 1	5	9	8	9	40	1	2	5
Professor 2	4	8	9	6	35	1	1	1
Professor 3	3	7	7	9	30	2	1	1
Professor 4	3	4	5	3	45	4	2	8
Professor 5	2	3	5	2	45	3	2	9
Professor 6	2	2	4	3	50	3	3	7

资料来源:取自 Taras 与 Kline(2007)。

研究者可以使用 Taras(2007)的 Dm-Index. Xls 计算该指标，所得结果如表 10-27。

表 10-27 Dm-index. xls 程序的输出结果

Correlations					
Hypothesized		Observed			
Commitment	Unwillingness to Move	Item 1	Item 2	Item 3	Item 4
0.5	−0.5	0.51	−0.04	0.22	0.14
0.5	−0.5	0.46	−0.23	0.54	0.26
0.5	−0.5	0.49	−0.45	0.56	0.17
0.5	−0.5	0.08	0.06	0.66	0.57
−0.5	0.5	−0.24	0.51	−0.75	−0.22
−0.5	0.5	−0.20	0.15	−0.69	−0.54
−0.5	0.5	−0.48	0.76	−0.60	0.11
−0.5	0.5	0.03	0.44	−0.65	−0.37

Dm for the total test	7.49	0.23	0.48	0.52
Dm for the 1st item deleted	6.86	0.29	0.53	0.47
Dm for the 2st item deleted	1.63	0.07	0.26	0.74
Dm for the 3st item deleted	7.25	0.30	0.55	0.45
Dm for the 4st item deleted	6.72	0.28	0.53	0.47

根据表 10-26 原始数据与表 10-27 底部的相关系数，研究者就可以整理出各题目与八个效标间的效度系数，请参见表 10-28 底部。

表 10-28 四个题目与八个效标间的理论与实得相关大小与相关方向

Exogenous Theoretically Relevant Variables	Correlations				
	Hypothesized direction of the relationsbips	Observed Correlations			
	Commiment	Item 1	Item 2	Item 3	Item 4
UR	0.5	0.51	−0.04	0.22	0.14
JS	0.5	0.46	−0.23	0.54	0.26
PS	0.5	0.49	−0.45	0.56	0.17
Coll	0.5	0.08	0.06	0.66	0.57
Age	−0.5	−0.24	0.51	−0.75	−0.22
Rel	−0.5	−0.2	0.15	−0.69	−0.54
SJ	−0.5	−0.48	0.76	−0.60	0.11
SJS	−0.5	0.03	0.44	−0.65	−0.37
删除某题后的 D_m 指标		0.47	0.74	0.45	0.47
总测量的 D_m 指标		0.52			

资料来源：修订自 Taras 与 Kline(2007)。

表 10-26 及表 10-28 中最左侧 UR 代表教授位阶,JS 代表工作满意度,PS 代表薪水满意度,COLL 代表合作机会,AGE 代表年龄,REL 代表亲戚人数,SJ 代表家庭成员的工作数,SJS 代表家庭成员的工作满意度。根据表 10-28 的理论与实得相关系数数值,研究者除了可以计算整体的效度系数(= 0.52)外,还可以计算某一题被删去后的效度系数以进行项目分析。删除某题的 D_m效度系数分别为:0.47,0.74,0.45,0.47。很明显,第二题关联效度不佳,可考虑删除,删除后整个测量的 D_m指标会显著提升(=0 .74)。

第七节　后　记

诚如王文中(1997)所言:研究者编制测量工具的主要目的在于“测量这个建构,而不是用工具来探索建构”。因此,不管是量表发展还是指标建构,理论建构应先于题目的编制,建构效度的检验应是验证性而非探索式。其次,就测量特性而言,原始分数通常并非线性,无法达到有意义测量(meaningful measurement)的基本要求:单向度、线性、不变性、客观性(Wright, 1999)。过去研究者运用多变量统计分析方法,例如:一般线性模型、因素分析、结构方程模型,均假定其依变量的测量具有线性与等距性。目前,研究者已可利用 Rasch 模型将原始分数转换为线性与等距性的测量,而使后续的统计分析及推论更具有效力。Rasch 分析的量尺,经丹麦数学家 Rasch(1960)证明不仅具有等距性,而且测量结果具有相对客观性:个人能力的估计不会随测量工具的不同而变动(item-free person measurement),题目难度的估计不会随不同受试群体而变动(person-free item calibration),测量编制者似乎可以善加应用,以奠定科学化的测量分析的基础。另外,由于题目的好坏取决于受试者是否认真作答,在做项目分析之前,应先利用 Rasch 分析软件(如 Winsteps,Bilog,Conquest)取得 personfit 指标,将异常作答的受试者剔除之后,再进行项目分析,才不至于将好的题目删除掉(因异常作答受试者存在而变成不良题目的情况)。当然受试者的 misfit 亦可能来自于不佳的题目,因此 item misfit 与 person misfit 的检查应是连续性的交替分析,一直到适配为止。由此观之,IRT(item response theory)理论在测量编制上的运用,似乎已开启了另一次测量统计的工业革命,量表编制的心理计量学将迈入崭新的境界。

本章习题

1. MacKenzie,Podsakoff 与 Jarvis(2005)认为转型领导内的指标,本质上是无法互换的,因此转型领导是一个组合型的潜在建构,如图 10-50 所示,而非如图 10-51 的原因建构。您同意吗? 为什么?

2. 图 10-52 是负向情绪的 CFA 模型,此测量模型界定正确吗? 为什么? 假如负向情绪是一个组合变量,要做如何限制才能使得该模型可辨识?

3. 图 10-53 是二阶的原因建构与效果建构的混合模型,模型中第一阶的薪水、同事、上司与升迁的满意度为效果建构,而第二阶的工作满意度则为原因建构,此种测量模型的界定可以成立吗?

4. 根据过去的文献,数常识(number sense)包含有五个因素,请利用图 10-54 的相关

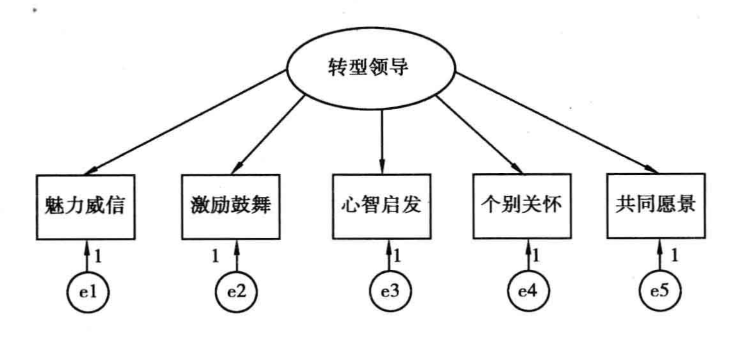


图 10-50 反映性的原因建构

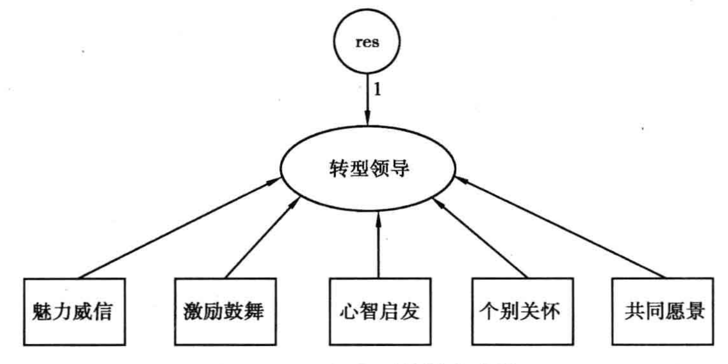


图 10-51 组合型的潜在建构

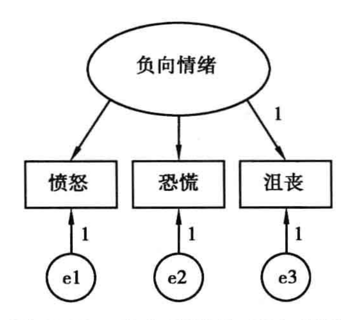


图 10-52 负向情绪的 CFA 模型

矩阵数据的信息,先绘制路径图,然后检验 Form1,Form2,Form3,Form5 的数常识测量的测量模型,以决定该测量模型的最佳模型,到底是平行测量还是同质性测量?接着利用 Amos 估计 Form1,Form2,Form3,Form5 的内部一致性信度,并与标准化的 Cronbach α 系数做比较。请特别注意 Form5 为四点量尺,其余为八点量尺。在同一测量中使用不同量尺,会影响测量的信度吗?

5. Cronbach α 并不是反映一个测量是否单向度的好指标,为何仍有很多的研究者依赖它来检验一个量表的内部同质性?探索式因素分析可补其不足吗?

6. 请根据图 10-55 中,Lewis,Templeton 与 Byrd(2005)的论述,评述量表发展的三阶段:范畴的界定、编拟题目、测量属性的评估,并说明相关的具体做法。

7. 探讨量表的效度检验,首推 Bagozzi(1981)的论文与 Straub,Boudreau ,Gefen 三人(2004)的经典论文,文中针对态度量表效度检验的过程与方法,做了具体而微的示范与说明,请详加研读并做心得摘要。

8. 为了检验提议模型是否界定错误(model misspecification),Glanville 与 Paxton(2007)利用验证性四分项协方差(CTA)分析,探究人与人之间的信任感,提出如图 10-56 中的两个信任感模型:心理习性模型与社会学习模型,心理习性模型是效果指标模型,社会学习模型乃是一种混合模型,包含效果指标与原因指标。验证性四分项协方差分析结果,摘要如表 10-29。

请根据表 10-29 的统计分析结果,回答下列问题:

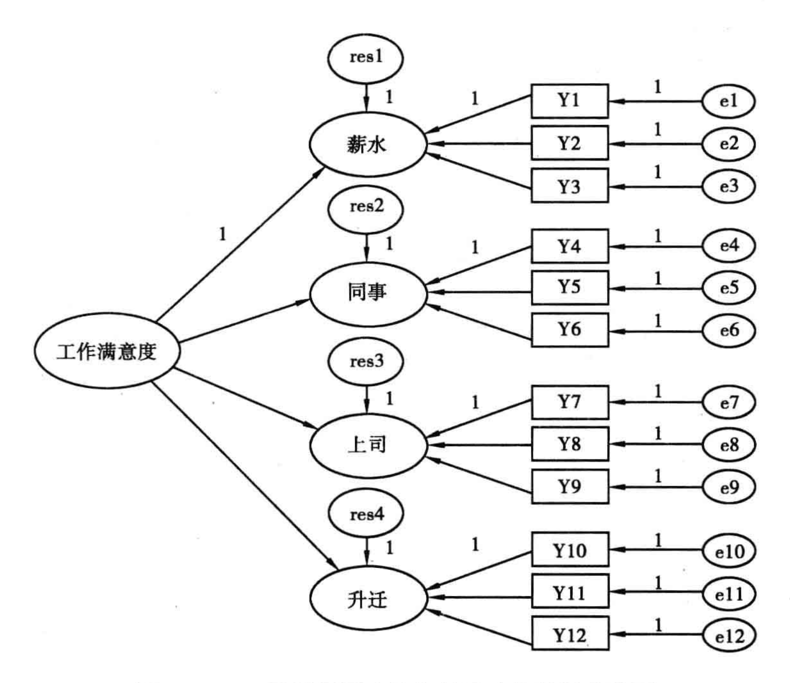


图 10-53　二阶的原因建构与效果建构的混合模型

alpha-sem-cor.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

15 : form4

	rowtype_	varname_	form1	form2	form3	form4	form5
1	n		1248.000	1248.000	1248.000	1248.000	1248.000
2	corr	form1	1.000	.160	.278	.243	.241
3	corr	form2	.160	1.000	.175	.161	.159
4	corr	form3	.278	.175	1.000	.195	.192
5	corr	form4	.243	.161	.195	1.000	.998
6	corr	form5	.241	.159	.192	.998	1.000
7	stddev		2.821	2.794	3.100	2.881	1.244
8	mean		3.890	4.190	4.060	4.250	1.830

图 10-54　五个数常识分测量的相关数据矩阵

(1)传统的 SEM 分析可以检验两个非嵌套设计的心理习性模型与社会学习模型吗?

(2)为何社会学习模型是心理习性模型的“tetrad-nested”模型?

Note:只有“tetrad-nested”的模型才能直接检验模型间的适配度差异。

(3)验证性 Tetrad 协方差分析的结果,显示出哪一个模型较适配?反映性测量模型或形成性测量模型?

(4)表 10-29 中的 nested test 统计量,旨在检验什么?所反映的意义是什么?

9. Helm(2005)曾以发展“公司声誉”(Corporate Reputation)量表为例,分四个步骤逐一说明如何编制一份形成性量表。该论文是一篇经典之作,请详加阅读后,说明这四个步骤如何具体实行?

10. Howell,Breivik 与 Wilcox(2007a)指出,因为形成性建构具有形成性指标的完全中介功能,它与各个外衍潜在变量的关系须具有一致性,亦即须具有外部一致性(external consistency)。以下将以他们论文中表 4 的数据为例(参见图 10-57),其中包含三个潜在变量:ses,eta2 与 eta3;八个观察变量,其中 x1 ~ x4 为形成性指标,y1 ~ y8 为反映性指标。

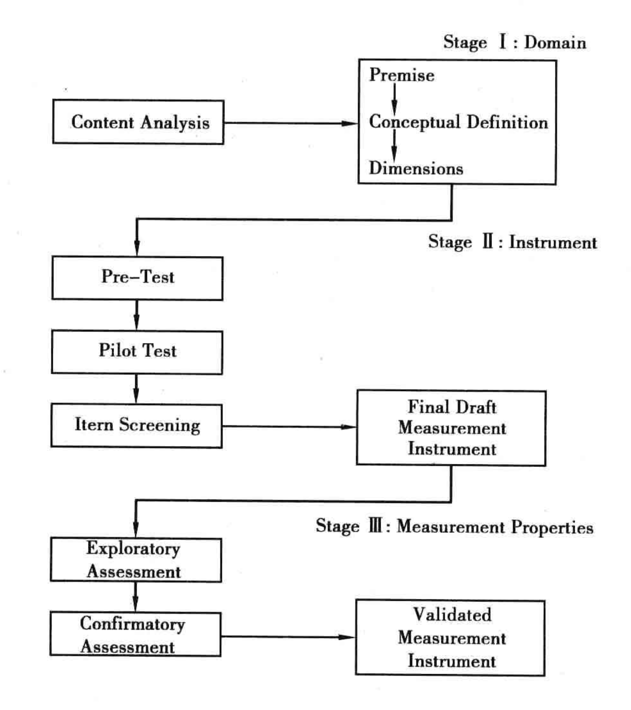


图 10-55 Lewis, Templeton 与 Byrd(2005)的量表发展三阶段

相关的路径图与分析结果如图 10-58。

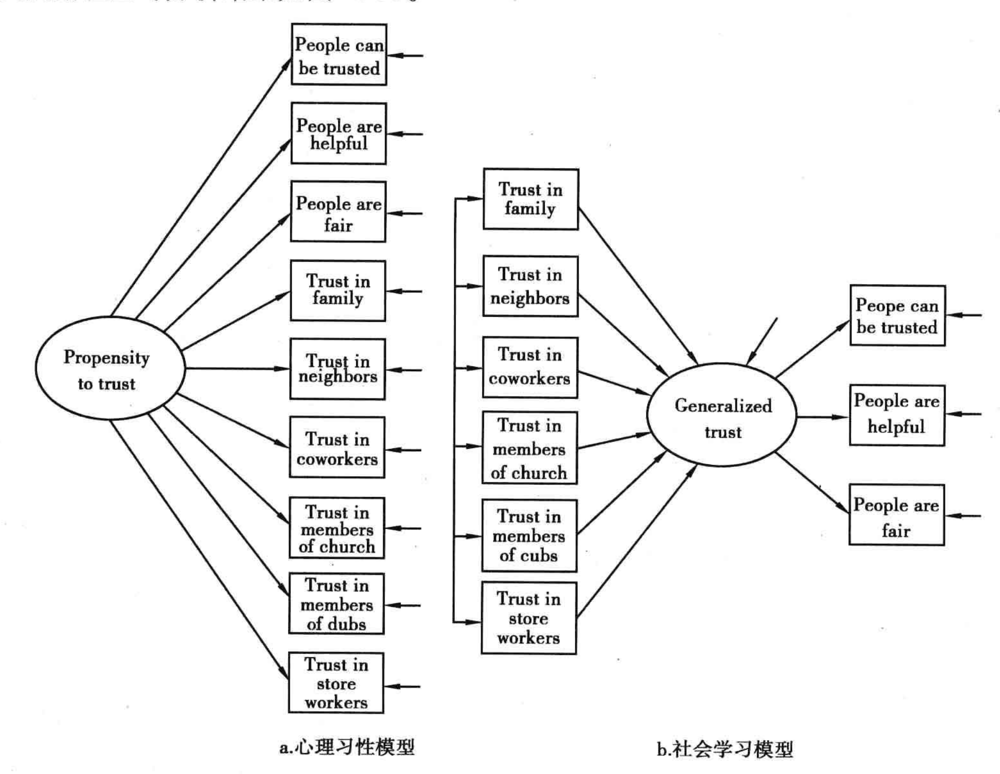


图 10-56 心理习性模型与社会学习模型

表 10-29　验证性 Tetrad 协方差分析摘要表

CTA, Listwise results	Psychological Propensiy Model			Social Learning Model			Nested Test		
	x^2	df	p-value	x^2	df	p-value	x^2	df	p-value
STS, National Sample (N=719)	73.55	27	0.000	8.59	12	0.737	64.95	15	0.000
STS, Phihdelphia Sample 1 (N=575)[2]	25.39	20	0.187	5.66	5	0.340	19.73	15	0.183
STS, Phihdelphia Sample 2 (N=584)[2]	33.47	20	0.029	3.33	5	0.650	30.15	15	0.011
SCBS (N=1927)[2]	16.89	5	0.005	NA	NA	NA	16.89	5	0.005

73.55−8.59 = 64.95　　27−12= 15

图 10-58 各变量间的相关系数显示,x1 与 x2 以及 y1 ~ y4(eta2)的关系显著高于 x3 与 x4 以及 y1 ~ y4(eta2)的关系,此种迹象可由图 10-59 的未标准化系数及表 10-30(A)的标准化系数反映出来,x3 与 x4 的未标准化系数与标准化系数显著低于 X1 与 X2。另外,值得注意的是,外衍潜在变量 eta2 与 eta3 所属的观察指标间的相关亦偏低(=0.25),显示出 eta2 与 eta3 的内部一致性不高(因为他们是 SES 的反映性建构)。理论上,他们之间的相关不可过低,否则会造成形成性建构失去单一解释性(unitary interpretability)的功能,而导致解释性的困惑(interpretational confounding)。换言之,SES 无法完全中介形成性指标与外衍建构(endogenous constructs,亦即 eta2 与 eta3)间的关系,而失去作为点变量(point variable)的功能,以致缺乏外部一致性(external consistency)。

File Edit View Data Transform Analyze Graphs Utilities Window Help

25 : y5

	rowtype_	varname_	x1	x2	x3	x4	y1	y2	y3	y4	y5	y6	y7	y8
1	n		432	432	432	432	432	432	432	432	432	432	432	432
2	corr	x1	1.00	.	.	.	.	.	.	.	.	.	.	.
3	corr	x2	.20	1.00	.	.	.	.	.	.	.	.	.	.
4	corr	x3	.20	.20	1.00	.	.	.	.	.	.	.	.	.
5	corr	x4	.20	.20	.20	1.00	.	.	.	.	.	.	.	.
6	corr	y1	.62	.59	.19	.09	1.00	.	.	.	.	.	.	.
7	corr	y2	.62	.59	.19	.09	.90	1.00	.	.	.	.	.	.
8	corr	y3	.62	.59	.19	.09	.90	.90	1.00	.	.	.	.	.
9	corr	y4	.62	.59	.19	.09	.90	.90	.90	1.00	.	.	.	.
10	corr	y5	.21	.26	.45	.39	.25	.25	.25	.25	1.00	.	.	.
11	corr	y6	.21	.26	.45	.39	.25	.25	.25	.25	.90	1.00	.	.
12	corr	y7	.21	.26	.45	.39	.25	.25	.25	.25	.90	.90	1.00	.
13	corr	y8	.21	.26	.45	.39	.25	.25	.25	.25	.90	.90	.90	1.00
14	stddev		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
15	mean		.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

图 10-57　SPSS 8 个变量的相关矩阵

表 10-30　形成性指标的标准化路径系数:摘录自 Amos 的输出结果

单向度 ses				双向度 ses			
ses	<---	x1	.6344	seC	<---	x1	.5564
ses	<---	x2	.5981	seC	<---	x2	.5165
ses	<---	x3	0.0463	seC	<---	x3	.5802
ses	<---	x4	−.1070	seC	<---	x4	.4679
eta2	<---	ses	.8792	eta2	<---	ses	.9907
eta3	<---	ses	.3159	eta3	<---	ses	.7040
Ⓐ				Ⓑ			

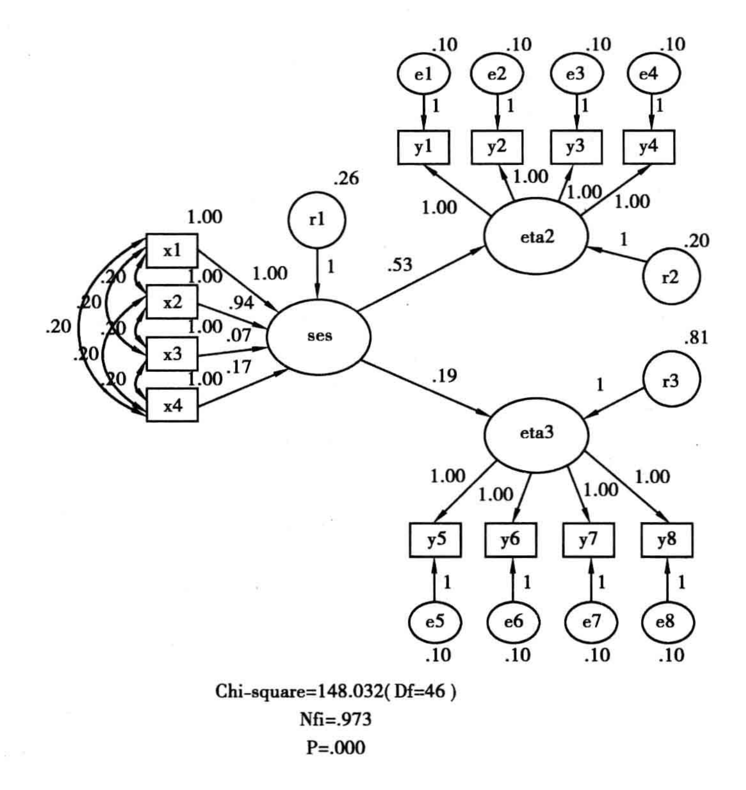


图 10-58　单向度形成性建构模型与分析结果

上述解释方面的困惑，Howell，Breivik 与 Wilcox（2007a）认为是出自于 SES 失去单一解释性功能所致，亦即 SES 可能含有两个不同的建构，笔者将此构念绘制如图 10-59 的双向度形成性建构模型，利用 Amos 进行数据分析，得到 χ^2 值 37.448，p 值为 0.864（df = 46），x1 ~ x4 的标准化系数如表 10-30（b）所示，显然其模型的适配度大为提升。根据前述的论证，请问形成性建构的向度及外衍建构间的关系，对于形成性测量有何重大影响？

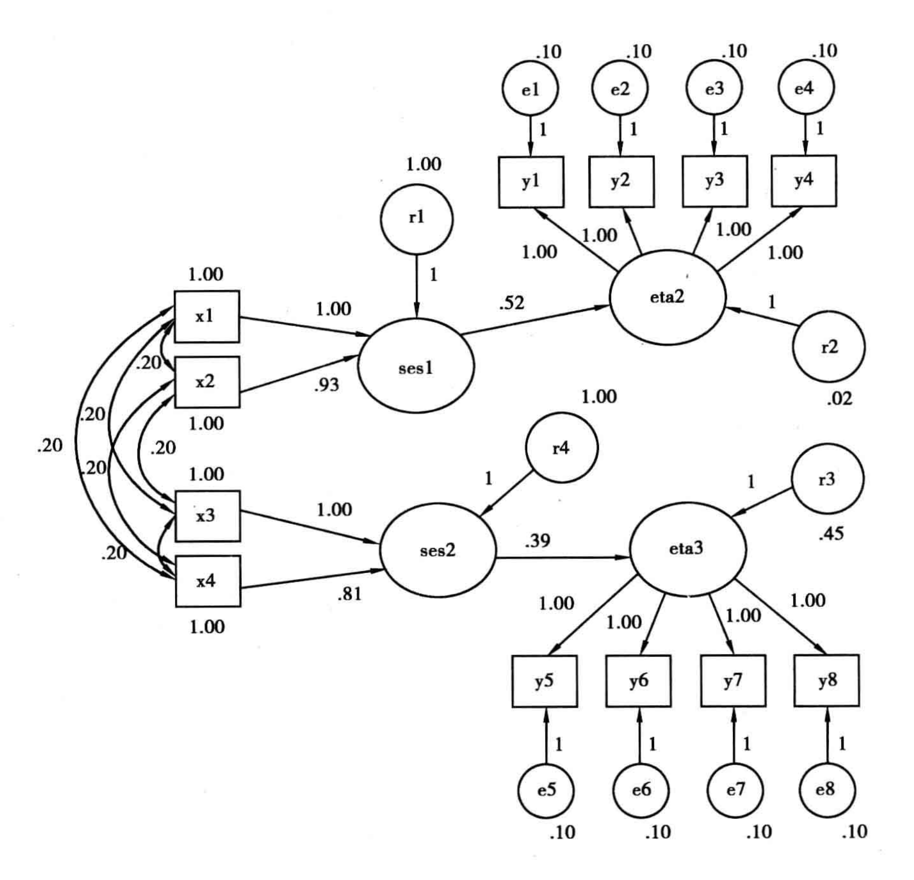


图 10-59　双向度形成性建构模型与分析结果

第十一章　结语：继往开来

Amos 从第 7 版之后，为了与 SPSS 的出版序号一致，版本序号从第 7 版直接跳至第 16 版。Amos 随着版次的增加，亦逐次增加了一些崭新功能，改版至今，以 6、7 版的变动较为显著，变动内容已在本书前三章中叙述，而第 16 版与第 17 版的新增功能，分别简略说明如下：

第一节　Amos 16.0 的新功能

Amos 16.0 新增了混合模型（Mixture modeling）的统计分析功能，混合模型适用于当理论模型不适用于总体，但适用于次群体时。例如，研究者发现所提议的回归模型不适用于总体，但却适用于次群体时，就可进行混合回归模型分析（Mixture regression modeling）。此外，混合模型亦可作为市场区隔研究（Market segmentation studies），取代 k-means 群聚分析，不用"距离"作为分类标准，而采用先验概率，计算个案属于某一特定团体的概率，将有共同特质的人或物归类在一块儿。Amos 可以利用事先归类的个案来分类机制，进行更有效的分群或归类分析。

第二节　Amos 17.0 的新功能

Amos 17.0 的主要新增功能：

1. 将路径图转换成 Visual Basic 程序，但反之则不可以；
2. 进行线性的成长曲线分析时，参数可以自动进行限制；
3. 路径图之间可以互相复制与剪贴。

由图 11-1 知，Amos 适时反映了当代统计学发展的整合轨迹（Grace，2008）：

此轨迹显示出传统统计与贝氏统计的结合（Lee，2007），Bayesian SEM 因而诞生，而 Amos 6.0 率先提供 Bayesian SEM 的统计分析。未来的统计方法亦可能从数据中"自我学习"，从数据中主动进行"区别分析"或搜寻"因果结构"。笔者深信，今后各版的 Amos 仍将承先启后，在整体的统计分析脉络上继往开来，追求卓越，成为更加智能的统计分析软件。

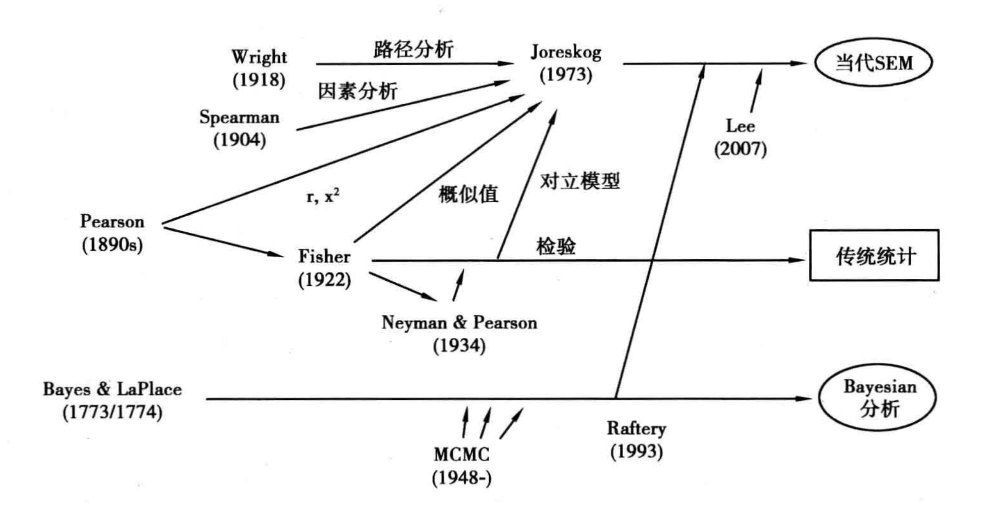


图 11-1 统计学发展轨迹

资料来源:http://www. structuralequations. com/resources/SEM + Essentials. pps

参考文献

中文部分

王文中(1997)。测量的建构:因素分析还是 Rasch 分析。调查研究,3,129-166。

余民宁(2006)。潜在变项模式:Simplis 的应用。台北:高等。

吴丽华(2007)。国小级任教师内外控信念与社会支持对身心健康影响之研究:以教师工作压力为中介变项。未出版的硕士论文,嘉义大学教育行政与政策发展研究所。

李茂能(1997)。评分者信度:Aiken 和谐系数之应用。嘉义师院学报,11,77-88。

李茂能(2006)。结构方程模式软件 Amos 之简介及其在测验编制上之运用:Graphics & Basic。台北:心理。

林姿饴(2005)。电脑化数常识量表之编制及其发展之研究:以九年一贯数学领域第二阶段学童为例。未出版硕士论文,嘉义大学数学教育研究所,嘉义。

邱皓政(2004)。结构方程模式:LISREL 的理论、技术与应用。台北:双叶。

陈智凯(2005)。知识经济指标之构念效度研究。长荣大学学报,9(1),105-134。

陈顺宇(2007)。结构方程模式:Amos 操作。台北:心理。

黄芳铭(2004)。社会科学统计方法学:结构方程模式。台北:五南。

黄芳铭,李俊贤(2008)。形成性与反映性潜在变项的对话。αβγ 量化研究期刊,2(1),69-97。

外文部分

Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, *40*, 955-959.

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory.* Belmont, California: Wadsworth.

Anderson, E. W. and Fornell, C. (2000). "Foundations of the American Customer Satisfaction Index," *Total Quality Management*, *11* (6), S869-S82.

Anderson, J. C., and Gerbing, W. (1988). "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach." *Psychological Bulletin*, 3:411-423.

Aspelmeier, J. (2007). Measurement Validity. Retrieved Nov. 12, 2007 from the World Wide Web: http://www.runet.edu/~jaspelme/611/Spring-2006/Measurment%20Validity.pdf

Bagozzi, R. P. (1981). An examination of the validity of two models of attitude. *Multivariate Behavioral research*, *16*, 323-359.

Bagozzi, R. P. (2007). On the meaning of formative measurement and how it differs from reflective measurement: Comment on Howell, Breivik, and Wilconx (2007). *Psychological Methods*, *12*, 229-237.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74-94.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling*, *9*(1), 78-102.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.

Berzruczko, N. (2004). Raw score nonlinearity obscures growth. *Rasch Measurement Transaction*, *18*(2), 973-974.

Biesanz, J. C., Deeb-Sossa, N., Papadakis, A. A., Bollen, K. A., & Curran, P. J. (2004). The Role of Coding Time in Estimating and Interpreting Growth Curve Models. *Psychological Methods*, *9*(1), 30-52.

Biesanz, J. C., Deeb-Sossa, N., Papadakis, A. A., Bollen, K. A., & Curran, P. J. (2004). The Role of Coding Time in Estimating and Interpreting Growth Curve Models. *Psychological Methods*, *9*(1), 30-52.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: John Wiley & Sons.

Bollen, K. A. (1993). *Liberal democracy: Validity and method factors in cross-national measures. American Journal of Political Sciences*, *37*(4), 1207-1230.

Bollen, K. A. (2007). Interpretational confounding is due to misspecification, not to type of indicator: Comment on Howell, Breivik, and Wilcox (2007). *Psychological Methods*, *12*, 219-228.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective.* New Jersey: John Wiley & Sons.

Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110*(2), 305-314.

Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110*(2), 305-314.

Bollen, K. A., & Ting, K. F. (2000). A Tetrad Test for Causal Indicators. *Psychological Methods*, *5*(1), 3-22.

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, *24*, 445-455.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological Methods and Research*, *5*, 3-51.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validity by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple Indicator latent growth modeling (MLGM). *Organizational Research Methods*, *1*, 421-483.

Cheung, M. W. L. (2007). Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models. *Structural Equation Modeling*, *14*, 227-246.

Chin, W. W. (1998). The Partial Least Squares Approach to Structural Equation Modeling, In *Modern Business Research Methods*, G. A. Marcoulides (ed.), Lawrence Erlbaum Associates, Mahwah, NJ.

Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, *16*, 64-73.

Churchill, G. A., Jr. (1995). *Marketing Research: Methodological Foundations*, 6th edition. Chicago: Dryden Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Coltman, T. R., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus Reflective Measurement Models: Two Applications of Erroneous Measurement. Journal of Business Research. Available online at www. sciencedirect. com.

Conway, J. M. (1998). Understanding method variance in multitrait-multirater performance appraisal matrices: Examples using general impressions and interpersonal affect as measured method factors. *Human Performance*, *11*, 29-55.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98-104.

Cote, J. A., & Buckley, R. (1987). Estimating trait, method, and error variance: Generalizing across 70 construct validation studies. *Journal of Marketing Research*, *24*, 315-318.

Crocker, L. M., Miller, M. D., & Franks, E. A. (1989). Quantitative methods for assessing the fit between test and curriculum. *Applied Measurement in Education*, *2*(2), 179-194.

Crocker, L., & Algina, J. (1986). *Introduction to classical & modern test theory.* New York: Holt, Rinehart & Winston.

Cronbach, L. J., & P. E. Meehl (1955). Construct Validity in Psychological Tests, *Psychological Bulletin*, *52*, 281-302.

Cudeck, R., & M. W. Browne. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research*, *18*,

147-167.

Curran, P. J. (1997). *Comparing three modern approaches to longitudinal data analysis: An examination of a single developmental sample.* Retrieved Jan. 1, 2009 from the World Wide Web: from http://www. unc. edu/? curran/ srcddocs/srcdmeth. pdf

Curran, P. J., & Hussong, A. M. (2003). The use of latent trajectory models in psychopathology research. *Journal of Abnormal Psychology*, *112*(4), 526-544.

Davis, L. (1992). Instrument review: Getting the most from your panel of experts. *Applied Nursing Research*, *5*, 104-107.

Diamantopoulos, A. (1999). Export performance measurement: Refiective versus formative indicators. *International Marketing Review*, *16*(6), 444-457.

Diamantopoulos, A., & H. Winklhofer (2001). Index construction with formative indicators: An alternative to scale development, *Journal of Marketing Research*, *37*, 269-277.

Diamantopoulos, A., & Siguaw, J. A. (2006). Formative Versus Reflective Indicators in Organizational Measure Development: A Comparison and Empirical Illustration. *British Journal of Management*, *17*, 263-282.

Doty, D. H., & Glick, W. H. (1998). Common method bias: Does common method variance really bias results? *Organizational Research Methods*, *1*, 374-406.

Driessen, G., Langen, A. van & Vierke, H. (2000). Basisonderwijs: veldwerkverslag, leerlinggegevens en oudervragenlijst. Basisrapportage PRIMA-cohortonderzoek. Derde meting 1998/99 [Primary education: research report, pupil's data and parent's questionary. Base report PRIMA cohort study. Third wave 1998/99]. Nijmegen: ITS, University of Nijmegen.

Duncan, S. C., & Duncan, T. E. (1994). Modeling incomplete longitudinal substance use data using latent variable growth curve methodology. *Multivariate Behavioral Research*, *29*, 313-338.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39-50.

Freedman, L. S., & Schatzkin, A. (1992). Sample size for studying intermediate endpoints within intervention trials of observational studies. *American Journal of Epidemiology*, *136*, 1148-1159.

Freeze, R. D., & Raschke, R. L. (2007). *An assessment of formative and reflective constructs in IS research.* Retrieved Feb. 12, 2008 from http://isz. lse. ac. uk/aspecis/20070055. pdf

Garson, G. D. (2008). Validity. Retrieved Feb. 22, 2008 from http://www2. chass. ncsu. edu/garson/pa765/validity. htm

Gelman, A. Carlin, J. B., Stern, H. S. & Rubin, D. B. (2004). *Bayesian Data Analysis.* (2nd ed.). Boca Raton: Chapman and Hall/CRC.

Gill, J. (2004). Introduction to the special issue. *Political Analysis*, *12*(4), 323-337.

Glanville, J. L., & Paxton, P. (2007). How do we learn to trust? A confirmatory tetrad analysis of the sources of generalized trust. *Social Psychology Quarterly*, *70*(3), 230-242.

Goo, J., Kim, D. J., & Cho, B. (2006). Structure of Service Level Agreements (SLA) in IT Outsourcing: The Construct and Its Measurement, 2006 Americas Conference on Information Systems (AMCIS), Acapulco, Mexico August 4-6.

Grace, J. (2008). SEM Essentials. Retrieved Dec. 12, 2008 from the World Wide Web: http://www. structuralequations. com/resources/SEM + Essentials. pps

Graham, J. M. (2006). Congeneric and (essentially) Tau-equivalent estimates of score Reliability: What they are and how to use them. *Educational and Psychological Measurement*, *66*, 930-944.

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in Nursing & Health*, *20*, 269-274.

Greene, M., & Way, N. (2005, June). Self-Esteem Trajectories among Ethnic Minority Adolescents: A Growth Curve Analysis of the Patterns and Predictors of Change. *Journal of Research on Adolescence* (*Blackwell Publishing Limited*), *15* (2), 151-177. Retrieved November 23, 2007 from Academic Search Premier database.

Gronlund, N. E. (1985). *Measurement and evaluation in teaching*(*5th ed.*). New York: MacMillan.

Gulliksen, H. (1950). *Theory of mental tests.* New York: Wiley.

Hadaya, P., Lége, P. M., Croteau, A. M. (2005). *Starting with the End in Mind: Building Capabilities to Integrate Electronic Commerce: Applications with Internal Information Systems.* HEC Montréal.

Hancock, G. R., Kuo, W., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural*

Equation Modeling: A Multidisciplinary Journal, *8*, 470-489.

Hartman, N. S., Williams, L. J., & Cavazotte, F. (2002). Common method variance: A demonstration of analysis with marker variables via partial correlations. *The Academy of Management 17*, *2*, *5*, *& 20-22*.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures 1*:77-89.

Helm, S. (2005). Designing a formative measure for corporate reputation. *Corporate Reputation Review*, *8*, 2, 95-109.

Hipp, J. R., Bauer, D. J., & Bollen, K. A. (2005). Conducting tetrad tests of model fit and contrasts of tetrad-nested models: A new SAS macro. *Structural Equation Modeling*, *12*(1), 76-93.

Hipp, J. R., Bauer, D. J., Curran, P. J, & Bollen, K. A. (2004). Crimes of opportunity or crimes of emotion: testing two explanations of season change in crime. *Social Forces*, *82*(4), 1333-1372.

Holbert, R. L., & Stephenson, M. T. (2002). Structural equation modeling in the communication sciences, 1995-2000. *Human Communication Research*, *28*,531-551.

Howell, R. D., Breivik, E., & Wilcox, J. B. (2007a). Reconsidering formative measurement. *Psychological Methods*, *12*, 205-218.

Howell, R. D., Breivik, E., & Wilcox, J. B. (2007b). Is Formative Measurement Really Measurement?

Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In Hoyle, R. H. (Ed.), *Structural equation modeling: Concepts, issues, and applications.* Thousand Oaks, CA: Sage.

Hulin, C., & Cudeck, R. (2001a). Cronbach's alpha on two-item scales. Journal of Consumer Psychology, *10*(1&2), 55.

Hulin, C., & Cudeck, R. (2001b). Can a reliability coefficient be too high? Journal of Consumer Psychology, *10*(1&2), 55-58.

Hulin, C., & Cudeck, R. (2001c). Why conduct a factor analysis and then compute an alpha? Journal of Consumer Psychology, *10*(1&2), 58-59.

Hulin, C., & Cudeck, R. (2001d). Why use alpha if it is not a good measure of unidimensionality? Journal of Consumer Psychology, *10*(1&2), 59-62.

Jackman, S. (2000). Estimation and inference via Bayesian simulation: An introduction to Markov chain Monte Carlo. *American Journal of Political Science*, *44*(2), 375-404.

James, L. R., Mulaik, S. A., & Brett, J. M. (2006). A Tale of Two Methods. *Organizational Research Methods*, *9*, 233-244.

Jarvis, C. B., Mackenzie, S. B., & Podsakoff, P. H. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, *30*, 199-218.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Distributions in Statistics: Continuous Univariate Distributions* (*Vol. 2*, 2nd ed.)., New York: Wiley.

Jonsson, F. Y. (1998), Modeling Interaction and Nonlinear Effects: A Step-by-Step Lisrel Example, in *Interaction and Nonlinear Effects in Structural Equation Modeling* (pp. 17-42), R. E. Schumaker and G. A. Marcoulides eds., Mahwah, NJ: Lawrence Erlbaum.

Joreskog, K. G. (1998). Interaction and nonlinear modeling: Issues and approaches. In R. E. Schumacker & G. A. Marcoulides (Eds), *Interaction and nonlinear effects in structural equation modeling.* New Jersey: Lawrence Erlabum Associates.

Kane, H. D., & Krenzer, D. (2006). A Confirmatory Analysis of the WAIS-Ⅲ: Using Data from Standardization and Independent Samples. *Counseling and Clinical Psychology Journal*, 2006, *3*(*3*), 113-136.

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, *96*(1), 201-210.

Kim, S., Murry, V., & Brody, G. (2001, April 1). *Studying the Relationship between Children's Self-Control and Academic Achievement: An Application of Second-Order Growth Curve Model Analysis.* (ERIC Document Reproduction Service No. ED452236)

Kline, R. B. (2004). *Principles and practice of structural equation modeling*(*2nd ed.*). New York: Guilford.

Law, K. S., & Wong, C. S. (1999). Multidimensional constructs in structural equation analysis: An illustration using the job perception and job satisfaction constructs. *Journal of Management*, *25*, 143-160.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563-567.

Lee, S. Y. (2007). *Structural equation modeling: A Bayesian approach.* Chichester, England: Wiley.

Lee, S. Y., & X. Y. Song. (2004). Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research*, *39*(4), 653-686.

Lewis, B. R., Snyder, C. A. & Rainer, K. R. Jr. (1995). An empirical assessment of the Information Resources Management construct. *Journal of Management Information Systems*, *12* (1), 199-223.

Lewis, B. R., Templeton, G. R., & Byrd, T. A. (2005). A methodology for construct development in MIS research. *European Journal of Information Systems*, *14*, 388-400.

Li, B., & Martin, E. B. (2002). An approximation to the F distribution using the chi-square distribution. *Computational Statistics & Data Analysis*, *40*,21-26.

Li, F. (1999). *Latent Curve Analysis: A Manual for Research Data Analysis.* Retrieved October 10, 2007 from the World Wide Web: http://osu. orst. edu/dept/hdfs/papers/paper. htm

Lindell, M. K., & Whitney, D. J. (2001). "Accounting for common method variance in cross-sectional research designs." *Journal of Applied Psychology*, *86*, 114-121.

Linhart, H., & W. Zucchini. (1986). *Model selection.* New York: John Wiley and Sons.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research*, *32*, 53-76.

Louis, V., & Jacqueline, H. (1997). A method for quantifying content validity of health-related questionnaires. *American Journal of Health Behavior*, *21*(1),67-72.

MacCallum, R. C., & Browne, M. W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin*,*114*(2), 533-541.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*, 490-504.

MacCallum, R. C., Roznowski, M., Mar, C. M., & Reith, J. V. (1994). Alternative strategies for cross-validation of covariance structure analysis. *Psychological Bulletin*, *114*, 185-199.

MacKenzie, S. B. and Spreng, R. A. (1992). How Does Motivation Moderate The Impact Of Central And Peripheral Processing On Brand Attitudes and Intentions? *Journal of Consumer Research*, *18* (March), 519-529.

MacKenzie, S. B., Podsakoff, P. M., & Jarvis, C. B. (2005). The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions. *Journal of Applied Psychology*, *90* (4), 710-730.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G. & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83-104.

MacKinnon, D. P., Lockwood, C., & Hoffman, J. (1998, June). *A new method to test for mediation.* Paper presented at the annual meeting of the Society for Prevention Research, Park City, UT.

MacKinnon, Fairchild, & Fritz (2007). Mediation Analysis. *The Annual Review of Psychology*, *58*, 593-614.

Malhotra, M. K., & Grover, V. (1998). An assessment of survey research in POM: from constructs to theory. *Journal of Operation Management*, *16*,407-425.

Manson, T. M., Levine, E. L., & Brannick, M. T. (2000). The Construct Validity of Task Inventory Ratings: A Multitrait-Multimethod Analysis. *Human Performance*, *13*(1), 1-22.

Marcoulides, G. A., & Schumacker, R. E. (1996). *Advanced structural equation modeling: Issues and techniques.* New Jersey: Lawrence Erlbaum Associates.

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 561-614). New York: Plenum.

Meade, A. W., Watson, A. M., & Kroustalis, C. M. (2007, April). *Assessing Common Methods Bias in Organizational Research.* Paper presented at the 22nd Annual Meeting of the Society for Industrial and Organizational Psychology, New York.

Mesquita, L. F., Anand, J., & Brush, T. (2005) Knowledge Transfer in Vertical Alliances: Advantages for Whom? *ACAC-Atlanta Conference for Competitive Advantage.* Atlanta, GA. June 24th.

Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling*, *2* (3), 255-273.

Muller, D. , Judd, C. M. , & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology*, *89*, 852-863.

Myung, J. I. , Karabatsos, G. , & Iverson, G. J. (2005). A Bayesian approach to testing decision making axioms. *Journal of Mathematical Psychology*, *49*, 205-225.

Patterson, G. R. (1993). Orderly change in a stable world: The antisocial trait as a chimera. *Journal of Consulting and Clinical Psychology*, *61*, 911-919.

Petter, S. , Straub, D. , & Rai, A. (2007). Specifying formative constructs in information systems research1. MIS Quarterly, *31*(4), pp. 623-665.

Ping, R. A. (1995). A Parsimonious estimating technique for interaction and quadratic latent variables. *Journal of Marketing Research*, *32* (August), 336-347.

Ping, R. A. (1996). Latent variable interaction and quadratic effect estimation A two-step technique using structural equation analysis. *Psychological Bulletin*, *119*, 166-175.

Ping, R. A. (2003). *Latent variable interactions and quadratics in survey data: A source book for theoretical model testing, 2nd edition.* [on-line monograph]. http://home. att. net/ ~ rpingjr/intquad2/toc2. htm.

Ping, R. A. (2005). "What is the average variance extracted for a latent variable interaction (or quadratic)?" [on-line paper]. Retrieved March 3, 2008 from the World Wide Web: http://home. att. net/ ~ rpingjr/ave1. doc)

Podsakoff, P. M. , MacKenzie, S. B. , Lee, J-Y, & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies, *Journal of Applied Psychology*, *88* (5), 879-903.

Preacher, K. J. & Leonardelli, G. J. (2006). *Calculation for the sobel test: An interactive calculation tool for mediation tests.* Retrieved June 5, 2007 from the World Wide Web: http://www. psych. ku. edu/preacher/sobel/sobel. htm

Preacher, K. J. , Rucker, D. D. , & Hayes, A. F. (2007). Addressing Moderated Mediation Hypotheses: Theory, Methods, and Prescriptions. *Multivariate Behavioral Research*, *42*(1), 185-227.

Raftery, A. (1993). *Bayesian model selection in structural equation models.* In K. Bollen & J. Long (Eds.), Testing structural equation.

Rasch, G. (1960). Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen: Danmarks Paedogogiske Institut, 1960 (Chicago: University of Chicago Press, 1980).

Raykov, T. , & Marcoulides, G. A. (2006). *A First Course in Structural Equation Modeling* (Second Edition). Mahwah, NJ: Lawrence Erlbaum Associates.

Reply to Bollen (2007) & Bagozzi (2007) *Psychological Methods*, *12*, 238-245.

Reuterber, S. E. , & Gustafsson, J. E. (1992). Confirmatory factor analysis and reliability: Testing measurement model assumptions. *Educational and Psychological Measurement*, *52*, 795-811.

Rindskopf, D. (1984). Using phantom and imaginary latent variables to parameterize constraints in linear structural models. *Psychometrika*, *49*, 37-47.

Schafer, J. L. (2000). *NORM* 2. 03 *for Windows 95/98/NT: Multiple imputation of incomplete multivariate data under a normal model.* http://www. stat. psu. edu/ ~ kls/misoftwa. html

Schafer, J. L. , &J. W. Graham. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147-177.

Schafer, J. L. , & M. K. Olsen. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research*, *33*(4), 545-571.

Scheines, R. , Hoijtink, H. , & Boomsma, A. (1999). Bayesian estimation and testing of structural equation models. *Psychometrika*, *64*, 37-52.

Schumacker, R. E. , & Marcoulides, G. A. (1998). *Interaction and nonlinear effects in structural equation modeling.* New Jersey: Lawrence Erlabum Associates.

Segars. A. H. , & Grover, V. (1998). Strategic information systems planning success: An investigation of the construct and its measurement. *MIS Quarterly*, *22*, 2, 139-163.

Shay, S. , & Gomez, J. (2002, April 1). *Privatization in Education: A Growth Curve Analysis of Achievement.* (ERIC Document Reproduction Service No. ED464115)

Shepard, L. A. (1993). Evaluating test validity. In L. Darling-Hammond, ed. , *Review of research in education*, *Vol. 19*,

pp. 405-450. Washington, DC: American Educational Research Association.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422-445.

Sireci, S. G. (1998). Gathering and analyzing content validity data. *Educational Assessment*, *5*(4), 299-311.

Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. In S. Leinhart (Ed.), *Sociological methodology* 1982 (pp. 290-312). San Francisco: Jossey-Bass.

Steiger, J. H. (1998). A note on multiple sample extensions of the MSEA fit index. *Structural Equation Modeling*, *5*, 411-419.

Stephenson, W. (1953). *The study of behavior: Q-technique and its methodology.* Chicago: University of Chicago Press.

Stoel, R. D., Wittenboer, G. Van den & Hox, J. (2004). Including time-invariant covariates in the latent growth curve model. *Structural Equation Modeling*, *11*(2), 155-167.

Straub, D., Boudreau, M-C., & Gefen, D. (2004). Validation guidelines for IS positive research. *Communications of AIS*, *13*, 380-427.

Taras, V., & Kline, T. (2007). *Quantifying Item Validity Using the Dh Index.* Paper presented at the Western Academy of Management Annual Conference, Missoula.

Tisak, J., & Meredith, W. (1990). Descriptive and associative development models. In A. Von Eye (Ed.), *Statistical methods in longitudinal research* (*Vol.* 2, pp. 387-406). Boston: Academic.

Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology*, *84*(3), 608-618.

Whittaker, T. A., & Stapleton, L. M. (2006). The performance of cross-validation indices used to select among competing covariance structure models under multivariate nonnormality conditions. *Multivariate Behavioral Research*, *41* (3), 295-335.

Willett, J. B. (1989). Questions and answers in the measurement of change. In E. Z. Rothkopf (Ed.), *Review of Research in Education*, *15*, 345-422. Washington, D. C.: AERA.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, *116*, 363-381.

Willett, J. B., & Sayer, A. G. (1996). Cross-domain analyses of change over time: Combing growth modeling and covariance structure analysis. In G. A. Marcoulides, & R. E. Schumackerm (Ed.), *Advanced structural equation modeling: Issues and techniques.* New Jersey: Lawrence Erlbaum Associates.

Willett, J. B., Ayoub, C. C., & Robinson, D. (1991). Using growth modeling to examine systematic differences in growth: An example of change in the functioning of families at risk of maladaptive parenting, child abuse, or neglect. *Journal of Consulting and Clinical Psychology*, *59*, 38-47.

Williams, L. J., Edwards, J. R., & Vandenberg, R. J. (2003). Recent Advances in Causal Modeling Methods for Organizational and Management Research. *Journal of Management*, *29*(6), 903-936.

Wilson, B., Callaghan, W., & Stainforth, G. (2007). An application of vanishing TETRAD analysis to a brand mode. *International Review of Business Research Papers*, *3* (2), 456-485.

Windle, M. (2000). A latent growth curve model of delinquent activity among adolescents. *Applied Developmental Science*, *4*, 193-207.

Windle, M. (2000). A Latent Growth Curve Model of Delinquent Activity Among Adolescents. *Applied Developmental Science*, *4*(4), 193-207.

Wright, B. D. (1999). *Fundamental measurement for psychology.* In S. E. Embretson & S. L. Hershberger (Eds.), The new rules of measurement: What every educator and psychologist should know (pp. 65-104 [http://www.rasch.org/memo64.htm]). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Wuensch, K. L. (2007). *Statistical Tests of Models That Include Mediating Variables.* Retrieved Dec. 5, 2007 from the World Wide Web: http://core.ecu.edu/psyc/wuenschk/MultReg/MediationModel.doc.

附录1　二层次因素分析VB.NET 程序：Amos 6.0 程序设计

```
#Region "Header"
Imports System
Import System. Diagnostics
Imports Microsoft. VisualBasic
Imports AmosEngineLib
#End Region
'2-level factor annalyis modelling by Fred Li, 2007
Module MainModule
  Public Sub Main( )
  Dim sem As New amosengine
Try
  SEM. TextOutput
  SEM. Ml
  SEM. Iterations (100)
  sem. Standardized
  'SEM. ModelMeansAndIntercepts
  SEM. BeginGroup ("C:\Program Files\amos 6\examples\JSP-BETWEN. SAV")
  SEM. GroupName ("BETWEEN GROUP MODEL")
  'C = 18.359 Obtained from MPLUS
  SEM. AStructure ("MATH1 = (4.285)MA1 + (1)NABILITY + E1(1)")
  SEM. AStructure ("MATH2 = (4.285)MA2 + (M2)NABILITY + E2(1)")
  SEM. AStructure ("MATH3 = (4.285)MA3 + (M3)NABILITY + E3(1)")
  SEM. AStructure ("ENG1 = (4.285)EN1 + (1)VABILITY + E4(1)")
  SEM. AStructure ("ENG2 = (4.285)EN2 + (EV2)VABILITY + E5(1)")
  SEM. AStructure ("ENG3 = (4.285)EN3 + (EV3)VABILITY + E6(1)")
  SEM. AStructure ("MA1 = (1)NABILITYB + (1)EB1")
  SEM. AStructure ("MA2 = NABILITYB + (1)EB2")
  SEM. AStructure ("MA3 = NABILITYB + (1)EB3")
  SEM. AStructure ("EN1 = (1)VABILITYB + (1)EB4")
  SEM. AStructure ("EN2 = VABILITYB + (1)EB5")
  SEM. AStructure ("EN3 = VABILITYB + (1)EB6")
  SEM. AStructure ("NABILITY(N1)")
  SEM. AStructure ("VABILITY(V1)")
  SEM. AStructure ("E1(ME1)")
  SEM. AStructure ("E2(ME2)")
```

```
  SEM. AStructure ("E3(ME3)")
  SEM. AStructure ("E4(EE1)")
  SEM. AStructure ("E5(EE2)")
  SEM. AStructure ("E6(EE3)")
  SEM. AStructure ("NABILITYB <-> VABILITYB")
  SEM. AStructure ("NABILITY <-> VABILITY(CO)")
  SEM. AStructure ("NABILITY <-> VABILITYB(0)")
  SEM. AStructure ("VABILITY <-> NABILITYB(0)")
  SEM. AStructure ("NABILITY <-> NABILITYB(0)")
  SEM. AStructure ("VABILITY <-> VABILITYB(0)")
  SEM. BeginGroup( "C:\Program Files\amos 6\examples\JSP-WITHIN. SAV")
  SEM. GroupName ("WITHIN GROUP MODEL")
  SEM. AStructure( "MATH1 = (1)NABILITY + E1(1)")
  SEM. AStructure ("MATH2 = (M2)NABILITY + E2(1)")
  SEM. AStructure ("MATH3 = (M3)NABILITY + E3(1)")
  SEM. AStructure ("ENG1 = (1)VABILITY + E4(1)")
  SEM. AStructure ("ENG2 = (EV2)VABILITY + E5(1)")
  SEM. AStructure ("ENG3 = (EV3)VABILITY + E6(1)")
  SEM. AStructure ("NABILITY(N1)")
  SEM. AStructure ("VABILITY(V1)")
  SEM. AStructure ("E1(ME1)")
  SEM. AStructure ("E2(ME2)")
  SEM. AStructure ("E3(ME3)")
  SEM. AStructure ("E4(EE1)")
  SEM. AStructure ("E5(EE2)")
  SEM. AStructure ("E6(EE3)")
  SEM. AStructure ("NABILITY <-> VABILITY(CO)")
  Debug. WriteLine("2-LEVEL FACTOR ANALYSIS" & vbcrlf & "Chi-square = " & FormatNumber (Sem. cmin,4))
  Debug. WriteLine("DF = "& Sem. df)
  Debug. WriteLine("P = "& FormatNumber(Sem. p,4) & vbcrlf & "Fred Li,2007")
Finally
  Sem. Dispose()
End Try
End Sub
End Module
```

为了避免与 VB. NET 的关键字相冲突,从 Amos 6.0 开始已将 SEM. Structure 改名为 SEM. Astructure,而且其后的模型设计或字串均须放在(　)之内,这是使用者须特别留神之处。

附录2 Number Sense 四因素修正结构：Amos 6.0 程序

```
#Region "Header"
Imports System. Math
Imports System. Diagnostics
Imports Microsoft. VisualBasic
Imports AmosEngineLib
Imports AmosGraphics
Imports AmosEngineLib. AmosEngine. TMatrixID
Imports PBayes
#End Region
Module MainModule
  Public Sub Main( )
  Dim Sem As New AmosEngine
  Try '除错开始
  Sem. TextOutput
  Sem. Standardized
  Sem. Mods (20)
  Sem. Smc
  sem. Corest
  'The 2 lines below used to compute SRMR
  Sem. NeedEstimates (14)
  Sem. NeedEstimates (11)
  Sem. BeginGroup ( "c:\NS94. SAV")
  Sem. AStructure ("q13 = (1)F1 + (1)e1")
  Sem. AStructure ("q10 = F1 + (1)e2")
  Sem. AStructure ("q9 = F1 + (1)e4")
  Sem. AStructure ("q14 = F1 + (1)e5")
  Sem. AStructure("q55 = (1)F2 + (1)e7")
  Sem. AStructure ("q45 = F2 + (1)e8")
  Sem. AStructure ("q50 = F2 + e(1)10")
  Sem. AStructure ("q43 = F2 + (1)e12")
  Sem. AStructure ("q29 = (1)F3 + (1)e13")
  Sem. AStructure ("q30 = F3 + (1)e14")
  Sem. AStructure ("q32 = F3 + (1)e15")
  Sem. AStructure ("q33 = F3 + (1)e16")
  Sem. AStructure ("q6 = (1)F4 + (1)e19")
```

```
Sem. AStructure ("q5 = F4 + (1)e20")
Sem. AStructure ("q4 = F4 + (1)e21")
Sem. AStructure ("q21 = F4 + (1)e22")
'MI for Model Modification
SEM. AStructure ("E4 <--> E20")
SEM. AStructure ("E2 <--> E12")
SEM. AStructure ("E2 <--> E13")
SEM. AStructure ("E5 <--> E22")
SEM. AStructure ("E8 <--> E22")
SEM. AStructure ("E16 <--> E19")
SEM. AStructure ("E4 <--> E7")
SEM. AStructure ("E5 <--> E21")
SEM. AStructure ("E1 <--> E21")
SEM. AStructure ("E1 <--> E7")
SEM. AStructure ("E5 <--> E12")
SEM. AStructure ("E5 <--> E13")
'2nd CFA
Sem. AStructure ("F1 = NS + (1)res1")
Sem. AStructure ("F2 = NS + (1)res2")
Sem. AStructure ("F3 = NS + (1)res3")
Sem. AStructure ("F4 = NS + (1)res4")
Sem. AStructure ("NS(1)")
"Below are the lines used to compute Standardized root mean squared residual(SRMR)
Dim N As Integer
Dim i As Integer
Dim j As Integer
Dim DTemp
Dim MESSAGE As String
Dim STATUS As Double
Dim MODELNAME As String
Dim Sample(,) As Double
Dim Implied(,) As Double
message = message & vbCrLf & vbCrLf & ModelName
  If Status <> 0 Then
  message = message & vbCrLf & " The model was not successfully fitted."
  Exit Sub
End If
  If Sem. AnyMissingValues Then
  message = message & vbCrLf & " The standardized RMR is not defined"
  message = message & vbCrLf & " when some data values are missing."
  Exit Sub
End If
Sem. GetEstimates (SampleCorrelations, sample)
Sem. GetEstimates (ImpliedCorrelations, implied)
N = UBound(Sample, 1)
'System. Diagnostics. Debug. WriteLine(N)
DTemp = 0
For i = 1 To N
```

```
    For j = 0 To i - 1
      DTemp = DTemp + (Sample(i,j) - Implied(i, j))^2
    Next
    Next
    N = N + 1'Array's index begibs with 0, so we have to add 1 to get the total N
    DTemp = Sqrt( DTemp / (N * (N + 1) / 2))
    'Dtemp is the standardized RMR
    Message = " Produced by Amos 6.0"
    If Status = 0 Then
    Message = Message & vbCrLf & " Standardized RMR = " & FormatNumber(DTemp,4)
    Else
    Message = Message & vbCrLf & " Sorry, the model was not successfully fitted. "
    End If
    System.Diagnostics.Debug.WriteLine (Message & vbCrLf & " Fred Li, 2006 嘉义大学")
    MsgBox (Message & vbCrLf & " Fred Li, 2006 嘉义大学"," 标准化 RMR")
    Finally
    SEM.Dispose()
    End Try
    End Sub
  End Module
```

附录3 间接效果检验的VB.NET 程序设计：Amos 6.0 程序

```
#Region "Header"
Imports System
Imports System. Diagnostics
Imports Microsoft. VisualBasic
Imports AmosEngineLib
Imports AmosGraphics
Imports AmosEngineLib. AmosEngine. TMatrixID
Imports PBayes
#End Region
Module MainModule
   'Example 6, Model C: Modified by Fred li, 2007
   'Wheaton et al. (1977).
      Sub Main( )
      Dim Sem As New AmosEngine
      Dim i As Integer
      Dim j As Integer
      Dim Ina As Double
      Dim Inb As Double
      Dim ind As Double
      Dim Indirect( , ) As Double
      Dim Indirectse( , ) As Double
      Dim Ztest( , ) As Double
      Dim Pvalue( , ) As Double
      Dim Addin As New Amosdebug. Amosdebug
      Try
         Sem. TextOutput( )
         Sem. Standardized( )
         Sem. Smc( )
         Sem. MonteCarlo( True)
         Sem. AllImpliedMoments ( )
         Sem. FactorScoreWeights ( )
         Sem. TotalEffects ( )
         Sem. NeedEstimates ( Indirecteffects)
         Sem. NeedStandarderrors ( Indirecteffects)
         Sem. BeginGroup ( AmosEngine. AmosDir & "Examples\Wheaton. sav" )
```

```
        Sem. AStructure ("anomia67 <--- 67_alienation (1)")
        Sem. AStructure ("anomia67 <--- eps1 (1)")
        Sem. AStructure ("powles67 <--- 67_alienation (path_p)")
        Sem. AStructure ("powles67 <--- eps2 (1)")
        Sem. AStructure ("anomia71 <--- 71_alienation (1)")
        Sem. AStructure ("anomia71 <--- eps3 (1)")
        Sem. AStructure ("powles71 <--- 71_alienation (path_p)")
        Sem. AStructure ("powles71 <--- eps4 (1)")
        Sem. AStructure ("67_alienation <--- ses(a)")
        Sem. AStructure ("67_alienation <--- zeta1 (1)")
        Sem. AStructure ("71_alienation <--- 67_alienation(b)")
        Sem. AStructure ("71_alienation <--- ses(c)")
        Sem. AStructure ("71_alienation <--- zeta2 (1)")
        Sem. AStructure ("educatio <--- ses (1)")
        Sem. AStructure ("educatio <--- delta1 (1)")
        Sem. AStructure ("SEI <--- ses")
        Sem. AStructure ("SEI <--- delta2 (1)")
        Sem. AStructure ("eps3 <--> eps1")
        Sem. AStructure ("eps1 (var_a)")
        Sem. AStructure ("eps2 (var_p)")
        Sem. AStructure ("eps3 (var_a)")
        Sem. AStructure ("eps4 (var_p)")
        Sem. FitModel ()
Sem. GetEstimates (Indirecteffects,indirect)
Sem. GetStandardErrors (Indirecteffects,indirectse)
'To get a specific parameter estimate
'ina = Sem. ParameterValue ("a")
'inb = Sem. ParameterValue ("b")
'ind = ina * inb
MsgBox ("Indirect Effect = " & FormatNumber( ind,4),," 间接效果")
Dim RNames () As String
Dim CNames () As String
Dim NR As Integer, NC As Integer

'Get the row and column variable names
Sem. RowNames (indirecteffects, RNames)
Sem. ColumnNames (indirecteffects, CNames)
'Print the standard errors
AdDin. Scientific ()
ADdin. FieldWidth = 15
ADdin. DecimalPlaces = 5
Debug. WriteLine (Vbcrlf & "Written by Fred Li, 2007, 嘉义大学")
Addin. PrintX (indirect,"间接效果")
Addin. PrintX (indirectse,"间接效果 SE")
System. diagnostics. Debug. WriteLine (vbCrLf & " Indirect Effects")
PrintMatrix (Indirect, CNames,RNames)
System. diagnostics. Debug. WriteLine (vbCrLf & "Standard errors for Indirect Effects")
PrintMatrix (IndirectSE, CNames,RNames)
NR = UBound(RNames)
```

```
NC = UBound(CNames)
Dim r As Integer
Dim c As Integer
Dim Ab As Double
ReDim ztest(nr,nc)
ReDim pvalue(nr,nc)
'To avoid Object reference not set to an instance of an object, the last line is required
For r = 0 To Nr
  For c = 0 To Nc
    If indirectse(r,c) =0 Then
    ztest(r,c) =0
    Else
    ztest(r,c) = indirect(r,c) / indirectse(r,c)
    If ztest(r,c) <0 Then ztest(r,c) = -(ztest(r,c))'取绝对值
    pvalue(r,c) = AmosEngine. ChiSquareProbability(ztest(r,c),1)
    End If
  Next
Next
System. diagnostics. Debug. WriteLine (vbCrLf & "间接效果 Z-test : CR")
PrintMatrix (Ztest, CNames,RNames)
System. diagnostics. Debug. WriteLine (vbCrLf & "P-values for 间接效果 Z-tests ")
PrintMatrix (Pvalue, CNames,RNames)
Sem. Dispose ()
    Finally
      Sem. Dispose()
    End Try
  End Sub
  'Print a matrix in the debug window
Sub PrintMatrix (ByVal TheMatrix(,) As Double, ByVal CNames $(), ByVal RNames $())
Dim NRows As Integer, NColumns As Integer
Dim i As Integer, j As Integer
NRows = UBound (RNames)
NColumns = UBound (CNames)
System. diagnostics. Debug. Write(" ")
For j = 0 To NColumns
System. diagnostics. Debug. Write (CNames(j). Padleft(20))
Next
System. diagnostics. Debug. WriteLine ("")
For i = 0 To NRows
System. Diagnostics. Debug. Write ( RNames(i). Padright (20))
For j = 0 To NColumns
System. Diagnostics. Debug. Write (TheMatrix(i, j). ToString(".0000"). PadLeft(10))
Next
System. Diagnostics. Debug. WriteLine ("")
Next
End Sub
End Module
```

附录4　建构信度的Amos 语法程序

```
#Region "Header"
Imports System. Math '要求数学函数运算用如(SQRT)
Imports System. Diagnostics
Imports Microsoft. VisualBasic
Imports AmosEngineLib
Imports AmosGraphics
Imports AmosEngineLib. AmosEngine. TMatrixID
Imports PBayes
#End Region
  Module MainModule
  Public Sub Main()
  Dim Sem As New AmosEngine
  Dim CNames() As String, RNames() As String
  Dim prt As New amosdebug. amosdebug '打印子串用
Sem. NeedEstimates (AllImpliedCovariances) '呼叫参数估计值用
Sem. NeedEstimates (SampleCovariances)
'Sem. InputUnbiasedMoments()
Sem. FitUnbiasedMoments()
  Try '除错开始
  Sem. TextOutput
  Sem. Standardized
  Sem. samplemoments
  sem. ResidualMoments
  Sem. Smc
  sem. Corest
  Sem. NeedEstimates (15)
  Sem. NeedEstimates (15)
  Sem. BeginGroup ( "F:\NS94. SAV")
  Sem. AStructure ("q4 = (a)F4 + (1)e1")
  Sem. AStructure("q5 = (b)F4 + (1)e2")
  Sem. AStructure("q6 = (c)F4 + (1)e3")
  Sem. AStructure("q21 = (1)F4 + (1)e4")
Dim N As Integer
Dim i As Integer
Dim j As Integer
Dim VarF(10) As Double
Dim VarE(10) As Double
```

```
Dim AllImplied( , ) As Double
Dim MESSAGE As String
Dim loadtot As Double
Dim vartot As Double
Dim Conrel As Double
Dim reg(10) As Double
Dim SReg As Double
Dim Sample( , ) As Double
reg(0) = 1
reg(1) = sem. ParameterValue ("c")
reg(2) = sem. ParameterValue ("b")
reg(3) = sem. ParameterValue ("a")
'必须与报表中的输出位置相呼应
System. Diagnostics. Debug. WriteLine (sem. ParameterValue ("a") & " a: q4 " ) '输出过程数据
System. Diagnostics. Debug. WriteLine (sem. ParameterValue ("b") & " b: q5 " )
System. Diagnostics. Debug. WriteLine (sem. ParameterValue ("c") & " c: q6 " )
For i = 0 To 3
prt. printx (FormatNumber (reg(i),4),"路径系数" & i)
Next
Sem. GetEstimates (AllImpliedCovariances, AllImplied)' 取得估计值
Sem. GetEstimates(SampleCovariances, Sample)
N = UBound(AllImplied)
  'Get regression weights needed for Construct reliability for the 1st factor
For i = 0 To 1
    For j = 0 To 1
  If i = j Then
    VarF(i) = AllImplied(i,j)
  'Get implied variance for latent trait
End If
    Next
Next
System. Diagnostics. Debug. WriteLine ( varf(0) & " Factor 4 Implied Var " )
Dim Evar As Double
Dim Sreg1 As Double
Dim SvarE(10) As Double
Dim path As Double
'累进标准化回归系数的计算
Sreg = 0:Evar = 0
  For i = 0 To 3
  For j = 0 To 3
  If i = j Then
    VarE(i) = AllImplied (i + 1, j + 1) - (reg(i) ^2 ) * VarF(0)'Get measurement errors
    prt. printx(FormatNumber(AllImplied(i + 1, j + 1),4), "指标方差" & i) '打印参数估计值
    prt. printx (FormatNumber(VarE(i),4), "测量误差变异量" & i)
    path = Sqrt (1 - VarE(i)/((reg(i) ^2 ) * VarF(0) + VarE(i )))
    prt. printx (FormatNumber(path,4),"标准化回归系数" & i)
    SReg = SReg + sqrt (1 - VarE(i)/((reg(i) ^2 ) * VarF(0) + VarE(i )))
    SVarE(i) = 1 - path ^ 2 'Get standardized variance for residuals
```

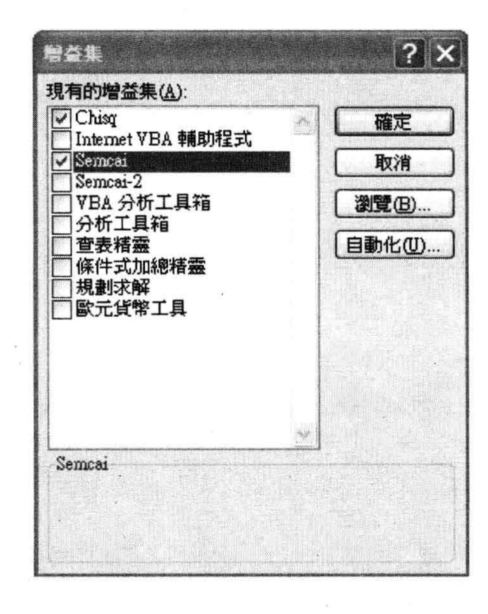
```
        Evar = Evar + SVarE(i)
        prt. printx (FormatNumber (sreg,4), "累进标准化回归系数" & i)
      End If
      Next
Next
  'System. Diagnostics. Debug. WriteLine (FormatNumber(Evar,4) & "累进测量误差和" )
  prt. printx (FormatNumber(Evar,4), "累进测量误差和")
  ConRel = Sreg ^ 2 /(Sreg ^ 2 + Evar)
  '计算建构信度估计值
  message = vbCrLf & FormatNumber(ConRel, 4)
  'System. Diagnostics. Debug. WriteLine( ConRel & "ConRel" )
  prt. printx (FormatNumber(ConRel, 4), "建构信度")
  MsgBox (Message & vbCrLf & " Fred Li, 2008 嘉义大学",, "建构信度" ) '要求显示讯息窗口
  Finally
  SEM. Dispose( )
  End Try
  End Sub
End Module
```

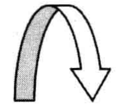
附录5 SEMCAI 增益集:SEMCAI.xla 的操作步骤

首先,将随书所附的软件 SEMCAI. xla 拷贝到 Microsoft 的 Addin 目录下(通常在:C:\Documents and Settings 次目录下),例如:



接着,再打开 Excel[工具]下的[增益集]选单,在所出现的窗口内点选 SEMCAI. xla 增益集,并按下[确定]。Excel 即会在主选单上出现[SEM-CAI]的选目(参看下图)。打开此选单会出现两个统计功能选目供您点选,第一个 CAI 为 SEM 样本规划,用以计算 SEM 研究的适切样本数。第二个 CAI 为 MI 统计分析,用以计算多元填补后的整合统计量。注意,输入数字前,请注意 Excel 的输入状态是否已设定在数字输入模型,这可免去切换成英文模型的困扰。





嘉義大學國教所, Fred Li, 2008 - Book1
檔案(F) 編輯(E) 檢視(V) 插入(I) 格式(O) 工具(T) SEM-CAI 統計(STAT) 資料
樣本規劃
MI統計分析
操作方法
離開

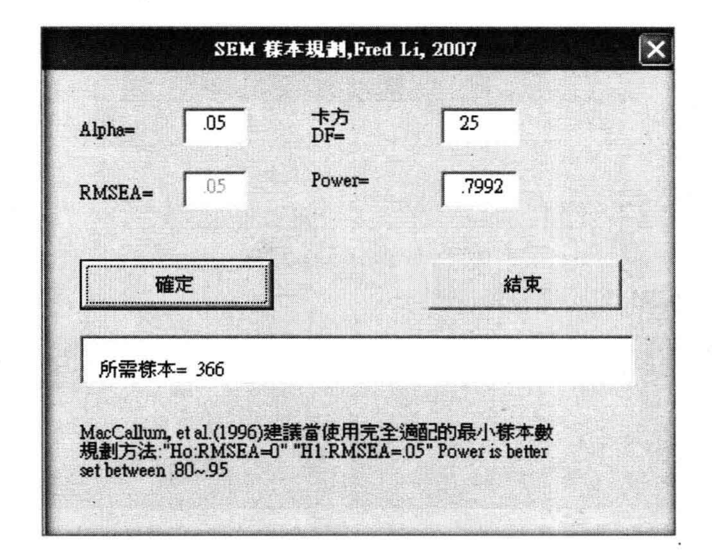
以下,逐一说明这两个 SEM 教学辅助组的实际操作步骤:

一、SEM 样本规划

1. 在以下 SEM 样本规划的窗口中输入 α、自由度及统计检验力。

SEM 樣本規劃,Fred Li, 2007
Alpha=
卡方 DF=
RMSEA= .05
Power=
確定
結束
MacCallum, et al.(1996)建議當使用完全適配的最小樣本數規劃方法:"Ho:RMSEA=0" "H1:RMSEA=.05" Power is better set between .80~.95

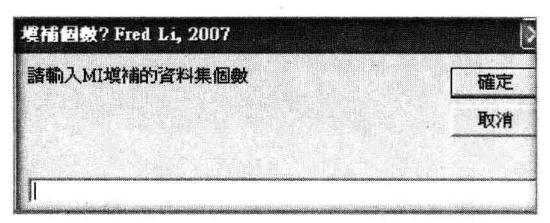
2. 执行结束后,会出现 SEM 研究所需的样本数。



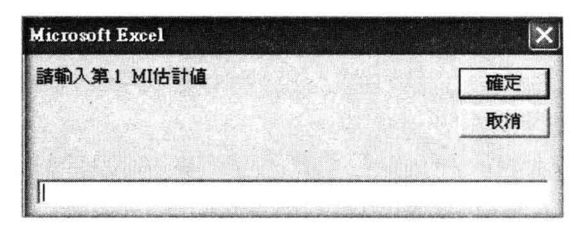
MI估计值	NU标准误
1.106	0.16
1.08	0.16
1.118	0.151
1.273	0.155
1.102	0.154
1.286	0.152
1.121	0.139
1.283	0.14
1.27	0.156
1.081	0.157

二、数据多元填补统计程序

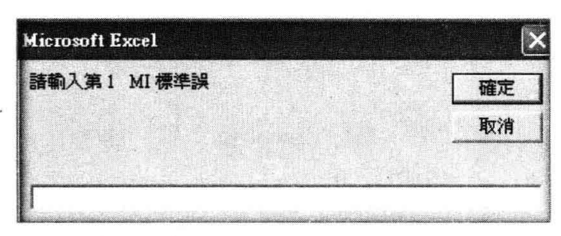
兹以下列数据集为例，说明多元填补（MI）分析副程序的操作步骤：



1. 输入 MI 多元填补的数据集个数



2. 输入单一参数的 MI 多元填补估计值



3. 输入单一参数的 MI 多元填补标准误

MI 平均數	t值	自由度	p值	臨界值
1.172	t=6.4869	df=109.2932	p=.0000	CR=1.982

MI 標準誤	.95 信賴區間
0.1807	Upper CI = 1.5301
	Low CI = .8139

4. 输出结果会在 Excel 表单上呈现

本程序所输出的 MI 多元填补统计，涵盖 MI 平均数、t 检验值、p 值及 0.95 的置信区间、MI 标准误、临界值等统计量。

附录6 Vanishing Tetrad 统计检验的步骤

一、Amos 的操作步骤

1. 如遇非等距以上变量时，请先利用 MPLUS 与 LISREL 等统计分析软件，取得多分相关协方差矩阵(polychoric covariance matrix)，以获得相关的协方差矩阵(asymptotic covariance matrix，简称 ACM)，目前 Amos 无法处理此类型的数据。

2. 利用 Amos 统计分析软件估计第一个提议模型，以取得样本协方差矩阵与隐含协方差矩阵。请先点选分析属性窗口[analysis properties]下的[output]，在出现的窗口中勾选[Sample moments]与[Impliedmoments]。执行完毕后，在输出窗口中，点选[Sample Covariances]以获得样本协方差矩阵，至于隐含协方差矩阵，则须点选[Implied Covariances]。

3. 欲进行嵌套模型检验(a nested test)时，须再利用 Amos 估计第二个提议模型，以取得隐含协方差矩阵，操作方法同前。

4. 拷贝前述 2 与 3 项的数据矩阵，整合进 CTANEST1. sas 程序中。

二、SAS CTANEST1 的操作步骤：SAS macro 实例解说

```
%include 'c:\ctanest1. mac';
%ctanest1(SAMPMAT1 =
1.44259240  0.32902571  0.24459959  0.18183673
0.32902571  1.53881889  0.40056103  0.29777936
0.24459959  0.40056103  1.58300286  0.43340765
0.18183673  0.29777936  0.43340765  1.60742121,
IMPMAT1B =
1.87624000  0.64992459  0.70311279  0.61217355
0.64992459  2.25867360  1.02575409  0.89308504
0.70311279  1.02575409  2.60323110  0.96617289
0.61217355  0.89308504  0.96617289  2.92354791,
IMPMAT2B =
 假如有第二个提议模型，请在此输入数据。
 …,
N =360, vars =4, nesttest =1, pchor =0, lisrel =0, mplus =0, lowdiag =0, reps =1);
  run;
```

上一行控制叙述中的选项设定须用“，”加以隔开。各项指令的用法说明如下：

SAMPMAT = 在此输入样本协方差矩阵(或多分相关矩阵)；各元素间须有空隔，亦

可以向量方式输入。

IMPMAT1B ＝ 在此输入第一个隐含协方差矩阵(或 vanishing tetrads 较多者,为一受限较多的模型)。

IMPMAT2B ＝ 在此输入第二个隐含协方差矩阵(或 vanishing tetrads 较少者,为一受限较少的模型)。

N ＝ 样本大小。

vars ＝ 观察变量数。

nesttest ＝ 当研究者想比较两个嵌套模型时,设定为1;假如仅有单一模型,设定为0。假如两个嵌套模型的比较未达显著差异,研究者会保留隐含四分项较多的模型。

pchor ＝ 当输入的多分相关矩阵为文件,其读取格式设定为 1,0 则为其他数据输入格式。

lisrel ＝ 当利用 LISREL 输入 ACM 相关矩阵时,设定为 1,0 则为其他输入格式。

mplus ＝ 当利用 MPLUS 输入 ACM 相关矩阵时,设定为 1,0 则为其他输入格式。

lowdiag ＝ 当输入的协方差矩阵或相关矩阵为下三角矩阵时,设定为 1,0 则为其他输入格式。

三、使用限制

本程序遇到自由度为 0 或卡方差异值为 0 时,将无法进行统计检验。其次,样本的大小亦要适中(请参阅附录 5),以免造成不当的统计检验力。另外,本程序尚无法正确检验非嵌套模型的统计分析;遇到此种状况,研究者须分别针对每一模型进行独立的分析(nesttest = 0),以获得正确的卡方值。之后,再进行卡方差异值的统计显著性检验。

※非嵌套模型统计分析的 SAS 报表

```
Model implied matrices are in correct order!!
ERROR: These models are NOT tetrad nested. Nested test is INAPPROPRIATE
          Tetrad Test for Model with more vanishing tetrads

               Chi-Square          df   p-value
               31.862452           14   0.0041903

          Tetrad Test for Model with fewer Vanishing tetrads

               Chi-Square          df   p-value
               22.113589           12   0.0362663

          Nested Tetrad Test for two models

               Chi-Square          df   p-value
```

术语对照表

A

Aaverage variance extracted　AVE 方法
Assumption of sphericity　球形假设
Autocorrelation　自变相关

B

Bayesian Estimation　贝氏估计法
Bayesian imputation　贝氏填补法
Bootstrap samples multiple subsamples of the same size
Burn-in　抽样分布未聚敛前所抛弃的样本数

C

Calibration sample　估计样本
Causal indicators　原因指标
Censored data　截断数据
Coding　编码
Common latent variable　共同潜在变量
Common method variance　共同方法变异
Component-based SEM　成分导向的 SEM
Composite latent variable　组合型潜在变量
Composite reliability　建构信度
Concurrent validity　同时效度
Conditional indirect effect　条件式中介/间接效果
Conditional latent change modeling　有条件的潜在特质变化的共变分析
Confidence level　信心水平/置信水平
Congeneric model　同质性测量模型
Constrained Model　限制模型
Construct reliability　建构信度
Construct validity　建构效度
Content relevance　领域关联性
Content specification　内容面向的界定
Content validity ratio　内容效度比
Content validity　内容效度
Convergence criterion　聚敛标准
Convergence in distribution　抽样分布的聚敛
Convergence of posterior summaries 事后摘要统计的聚敛
Convergent validity　聚敛效度
Correlated-trait correlated-method model　特质相关—方法相关模型
Correlated-trait correlated-uniqueness model　特质相关—误差相关模型
Covariance-based SEM 协方差导向的 SEM
Criterion validity　效标关联效度
Critical ratio　临界比
Critical Ratios for Differences Between Parameters 参数间的差异性检验
Cross-domain analysis of change over time　纵贯式跨领域研究
Cross-validation index　复核效度指标
Cross-validation　复核效度
CS　Convergence satatistics
CS　聚敛统计量
Custom Estimation　自订参数的估计

D

Data imputation　缺失数据填补
Decompose the formative construct　形成性指标的建构加以解组
DIC Deviance information criterion
Different antecedents and consequences　不同的前因与后果
Diffuse prior distribution　均分性的先验概率分布
Dimensionality　潜在变量的向度
Discriminant validity　区辨效度
Domain definition　领域定义
Domain representation　领域代表性
Double cross-validation　效度双复核
Dummy Coding　虚拟编码
Dummy/Effect variables　虚拟/效果变量

E

ECVI Expected cross-validation index
Effect Coding 效果值编码
Effect construct 效果建构
Effect indicators 效果指标
Endogenous variables 内衍变量
Essential 必要
Expectancy-value measures 预期价值法
Exponential function 指数函数
External consistency 外部一致性
External validity 外在效度

F

Factorial validity 因素效度/因素结构效度
Factor-of-curves model 曲线因子模型
Formative indicator model 形成性指标测量模型
Formative indicators 形成性指标

G

Gain score or change score 平均数差异值
Generalizability theory 推论力理论
Global item 共通性题目
Growth curve model 成长曲线模型

H

Hbrid model 混合模型
Histogram 条图

I

Implied correlations 隐含相关矩阵
Implied sample moment 隐涵样本动差
Improper solution 不适当解
Index construction 指标建构
Index 指标
Indicant product approach 指标交乘积直接估计法
Indicator collinearity 指标间共线性的检验
Indicator specification 观察指标的界定
Interact with the effect of an IV on the outcome DV 因调节变量而产生交互作用现象
Internal consistency 内部一致性
Interpretational confounding 解释性困惑
Interrater reliability 评分者间信度
Intraindividual differences in change 个体内潜在特质变化的情形
IRT Item response theory 项目反应理论
Item construct validity 题目建构效度
Item parceling 题组分析法
Item reliability 题目的信度指标
Item validity 题目效度
Item-free person measurement 个人能力的估计不会随测量工具的不同而变动
Item-objective congruence index 题目—目标一致性指标

K

Key shortcuts 快速键

L

Label 标签
Latent change analysis 潜在特质变化分析
Latent growth models 潜在特质成长分析
Latent variable longitudinal curve model 潜在变量纵贯式曲线模型
Level factor 测量基线因子
Logic check 逻辑比较分析
Loose replication strategy 宽松复核取向

M

Marker variable 标竿指标
Market segmentation studies 市场区隔研究
Markov Chain Monte Carlo 马克夫链模拟法
Max observations 最大样本保留数
Mean-centering 离均差分数
Meaningful measurement 有意义测量
Measure of generalizability 推论力的指标
Measurement invariance across time or longitudinal measurement invariance 纵贯性测量不变性的假设
Measurement model misspecification 测量模型的界定错误
Mediating variables 中介变量
Method factor 方法因子
Mixture modeling 混合模型
Mixture regression modeling 混合回归模型
Mmodel misspecification 模型界定错误
Model Identification 模型是否可辨识
Model selection 模型选择性
Model stability 模型稳定性
Moderate replication strategy 适中复核取向
Moderated mediation 调节型(或称条件化)中介效果
Moderation variables 调节变量
Multicollinearity 多元共线性
Multidimensional construct 多向度的建构
Multiple imputation 多元填补技术
Multiple indicator latent growth model 多重指标潜在特质成长模型
Multiple indicators and multiple causes of a single latent trait MIMIC 模型
Multiple product indicator approach 多元指标法
Multiple-imputation estimate of the regression weight

多元填补回归系数的估计值
Multitrait-multimethod approach 多项特质—多项方法/MTMM 方法

N

NCDS National Child Development Study
Nested design 嵌套/巢式设计
Nomological Validity 构念网络效度
Non-informative Prior distribution 均分性的先验概率分布
Non-nested models 非嵌套的模型
Non-numeric data 非数据数据
Not necessary 没有必要
Number of batches for batch means 样本批组数
Number of burn-in observations 抛弃样本数
Number of estimands 估计参数个数
Number of imputations 填补数据集的个数
Number sense 数常识

O

Ordered-categorical 次序性类别数据
Orthogonal Coding 正交编码

P

Perfect convergence 完美聚敛
Person-free item calibration 题目难度的估计不会随不同受试群体而变动
Phantom variable 幽灵变量
Piecewise growth modeling 片段线性成长模型
Ping's 2-step Ping's 双步骤
Plugins 增益集
Point variable 点变量
Polygon 多边图
Population moment 总体动差
Posterior distribution 事后概率分布
Posterior predictive distribution 事后预测概率分布
Posterior simulation 事后模拟
Predictive validity 预测效度
Prescriptive validity 处方效度
Prior distribution 估计参数的先前信念
Prototypic moderated mediation 典型调节型中介效果

Q

Quadratic function 二次函数
Quadratic polynomial 二次多项式

R

Random numbers 随机数
Reflective indicator model 反映性指标测量模型
Reflective indicators 反映性指标
Regression imputation 回归填补法
Replications 复制研究
Residual direct effect 净直接效果

S

Scale development 量表发展
Scale 量表
Second-order growth modeling 二阶成长模型
Self-report、peer-report & diary behavior report 自我报告、同侪报告法与日记
Semantic differential measures 语意区分法
Shaded 阴影
Shape factor 改变形态因子
Single indicator approach 单指标法
Spurious relationship 虚假相关
Stochastic regression imputation 随机回归填补法
Suppressor variable 抑制变量

T

Target pdf 目标条件概率密度函数
Thinning 数据瘦身
Tight replication strategy 严谨复核取向
Tuning parameter MCMC 的微调参数

U

Unconstrained Model 开放模型
Unidimensionality 单一向度
Uniform 等概率
Unitary concept 单一的概念
Unitary interpretability 单一解释性
Useful but not essential 有用但不必要

V

Validation sample 效度样本
Validity extension 进行效度延伸性
Validity generalization 效度类推性
Validity 效度
Vanishing tetrad 消失四分项
VIF variance inflation factor 方差膨胀因子

万卷方法总书目

万卷方法是我国第一套系统介绍社会科学研究方法的大型丛书，来自中国社科院、北京大学等研究机构和高校的两百余名学者参与了丛书的写作和翻译工作。至今已出版图书60多个品种，其中绝大多数是2007年以来出版的新书。

1 图解 AMOS 在学术研究中的应用
978-7-5624-6223-1

2 应用 STATA 做统计分析(更新至 STATA10.0)
978-7-5624-5986-6

3 社会调查设计与数据分析——从立题到发表
978-7-5624-6074-9

4 质性研究导引
978-7-5624-6132-6

5 APA 格式——国际社会科学学术写作规范手册
978-7-5624-6105-0

6 如何做心理学实验
978-7-5624-6151-7

7 话语分析导论
978-7-5624-6075-6

8 心理学学位论文写作全程指导
978-7-5624-6113-5

9 心理学研究方法导论
978-7-5624-5828-9

10 分类数据分析
978-7-5624-6133-3

11 结构方程模型:AMOS 的操作与应用(附光盘版)
978-7-5624-5720-6

12 AMOS 与研究方法(第 2 版)
978-7-5624-5569-1

13 爱上统计学 (第 2 版)
978-7-5624-5891-3

14 社会科学定量研究的变量类型、方法选择与范例解析
978-7-5624-5714-5

15 案例研究:设计与方法(中译第 2 版)
978-7-5624-5732-9

16 问卷设计手册:市场研究、民意调查、社会调查、健康调查指南
978-7-5624-5597-4

17 广义潜变量模型:多层次、纵贯性以及结构方程模型
978-7-5624-5393-2

18 调查问卷的设计与评估
978-7-5624-5153-2

19 心理学论文写作——基于 APA 格式的指导
978-7-5624-5354-3

20 心理学质性资料的分析
978-7-5624-5363-5

21 问卷统计分析实务:SPSS 操作与应用
978-7-5624-5088-7

22 如何做综述性研究
978-7-5624-5375-8

23 质性访谈方法
978-7-5624-5307-9

24 量表编制:理论与应用(校订新译本)
978-7-5624-5285-0

25 质性研究:反思与评论(第 2 卷)
978-7-5624-5143-3

26 实验设计原理:社会科学理论验证的一种路径
978-7-5624-5187-7

27 混合方法论:定性研究与定量研究的结合
978-7-5624-5110-5

28 社会统计学
978-7-5624-5253-9

29 校长办公室的那个人(质性研究个案阅读)
978-7-5624-4880-8

30 泰利的街角(质性研究个案阅读)
978-7-5624-4937-9

31 客厅即工厂(质性研究个案阅读)
978-7-5624-4886-0

32 标准化调查访问
978-7-5624-5062-7

33 解释互动论
978-7-5624-4936-2

34 如何撰写研究计划书
978-7-5624-5087-0

35 质性研究的理论视角:一种反身性的方法论
978-7-5624-4889-1

36 社会评估:过程、方法与技术
978-7-5624-4975-1

37 如何解读统计图表
978-7-5624-4906-5

38 公共管理定量分析:方法与技术(第 2 版)
978-7-5624-3640-9

39 量化研究与统计方法
978-7-5624-4821-1

40 心理学研究要义
978-7-5624-5098-6

41 调查研究方法 (校订新译本)
978-7-5624-3289-0

42 分析社会情境:质性观察和分析方法
978-7-5624-4690-3

43 建构扎根理论:质性研究实践指南

978-7-5624-4747-4

44 参与观察法
978-7-5624-4616-3

45 文化研究:民族志方法与生活文化
978-7-5624-4698-9

46 质性研究方法:健康及相关专业研究指南
978-7-5624-4720-7

47 如何做质性研究
978-7-5624-4697-2

48 质性研究中的访谈:教育及社会科学研究者指南
978-7-5624-4679-8

49 案例研究方法的应用(中译第 2 版)
978-7-5624-3278-3

50 教育研究方法论探索
978-7-5624-4649-1

51 实用抽样方法
978-7-5624-4487-9

52 质性研究:反思与评论(第 1 卷)
978-7-5624-4462-6

53 社会科学研究的思维要素(第 8 版)
978-7-5624-4465-7

54 哲学史方法论十四讲
978-7-5624-4446-6

55 社会研究方法
978-7-5624-4456-5

56 质性资料的分析:方法与实践(第 2 版)
978-7-5624-4426-8

57 实用数据再分析法(第 2 版)
978-7-5624-4296-7

58 质性研究的伦理
978-7-5624-4304-9

59 叙事研究:阅读、倾听与理解
978-7-5624-4303-2

60 质化方法在教育研究中的应用(第 2 版)
978-7-5624-4349-0

61 复杂调查设计与分析的实用方法(第 2 版)
978-7-5624-4290-5

62 研究设计与写作指导:定性、定量与混合研究的路径
978-7-5624-3644-7

63 做自然主义研究:方法指南
978-7-5624-4259-2

64 多层次模型分析导论(第 2 版)
978-7-5624-4060-4

65 评估:方法与技术(第 7 版)
978-7-5624-3994-3

66 焦点团体:应用研究实践指南(第 3 版)
978-7-5624-3990-5

67 质的研究的设计:一种互动的取向 (第 2 版)
978-7-5624-3971-4

68 组织诊断:方法、模型和过程(第 3 版)
978-7-5624-3055-1

69 民族志:步步深入(第 2 版)
978-7-5624-3996-7

70 分组比较的统计分析(第 2 版)
978-7-5624-3942-4

71 抽样调查设计导论 (第 2 版)
978-7-5624-3943-1

72 定性研究(第 1 卷):方法论基础(第 2 版)
978-7-5624-3851-9

73 定性研究(第 2 卷):策略与艺术 (第 2 版)
978-7-5624-3286-9

74 定性研究(第 3 卷):经验资料收集与分析的方法(2 版)
978-7-5624-3944-8

75 定性研究(第 4 卷):解释、评估与描述(第 2 版)
978-7-5624-3948-6

76 社会网络分析法 (第 2 版)
978-7-5624-2147-4

77 公共政策内容分析方法:
978-7-5624-3850-2

78 复杂性科学的方法论研究
978-7-5624-3825-0

79 社会科学研究:方法评论
978-7-5624-3689-8

80 论教育科学:基于文化哲学的批判与建构
978-7-5624-3641-6

81 科学决策方法:从社会科学研究到政策分析
7-5624-3669-0

82 电话调查方法:抽样、筛选与监控 (第 2 版)
7-5624-3441-7

83 研究设计与社会测量导引(第 6 版)
978-7-5624-3295-1

为了建设好“万卷方法”,更好地服务学界,现由重庆大学出版社和人大经济论坛做出决定,凡购买重庆大学出版社的万卷方法系列图书的读者,填写以下信息调查表(复印即可),邮寄给我们(400030　重庆大学出版社　林佳木),经过认证后,我们将会赠送人大经济论坛币 100 个(可免费下载丛书相关学习资料并与教师及学友进行交流):

读者情况调查表	
姓名	
单位	
联系电话	
E-mail	
论坛 ID	
使用书籍	
购买渠道	
对丛书建设的建议	
邮政地址(邮编)	

人大经济论坛

——国内最大的经济、管理、金融、统计类在线教育网站

人大经济论坛(网址:http://www.pinggu.org)依托中国人民大学经济学院,于 2003 年成立,致力于推动经济学科的进步,传播优秀教育资源,目前已经发展成为国内最大的经济、管理、金融、统计类的在线教育和咨询网站,也是国内最活跃和最具影响力的经济类网站。

1. 拥有国内经济类教育网站最多的关注人数,注册用户以百万计,日均数十万经济相关人士访问本站。

2. 是国内最丰富的经管类教育资源共享数据库和发布平台。

3. 论坛给所有会员提供学术交流与讨论的平台,同时也有网络社交 SNS 的空间,经管百科提供了丰富专业的经管类在线词典,数据定制和数据处理分析服务是您做实证研究的好帮手,免费的经济金融数据库使您不再为数据发愁,更有完善的经管统计类培训和教学相关软件,只要您是学习、研究或从事经管类行业,人大经济论坛就能满足您的需要!